Prefaţă

Volumul de faţă reprezintă rodul unei activităţi susţinute şi relativ îndelungate a celor doi autori, în domeniul modelării ce are ca scop evaluarea calităţii servirii în reţelele de telecomunicaţii, calitate ce este definită, conform cu normativele ITU-T, avizul E.800, ca fiind *"efectul rezultant al performanţelor unui serviciu care determină gradul de satisfacţie al utilizatorului o dată cu folosirea respectivului serviciu"*. Se subliniază astfel faptul că o combinaţie de aspecte caracterizează QoS, şi anume: suportul şi operativitatea serviciului, precum şi accesibilitatea şi integritate sa.

Concret, lucrarea, dat fiind complexitatea subiectului abordat, este alcătuită din trei părţi cu un oarecare grad de autonomie, care în ansamblu oferă cadrul necesar îndeplinirii demersului propus, cel al studiului cantitativ legat de calitatea serviciilor de telecomunicaţii. Astfel, după o introducere necesară alcătuirii unei imagini de ansamblu a subiectului abordat, prima parte aduce în discuţie concepte de ordin teoretic şi practic, ce ţin de bagajul fundamental de cunoştinţe din domeniu. Trei capitole sunt dedicate acestui scop, fiecare prezentând, pe rând, *factorii* cu influenţă determinată asupra calităţii, *mijloacele* uzuale de management al calităţii *şi metodele* de asigurare a calităţii, specifice reţelelor IP

Odată trecute în revistă principalele aspecte de natură practică, implicate în controlul calităţii, urmează a doua parte a lucrării, în care se prezintă instrumentul matematic ce stă la baza activităţilor descrise în a treia parte, de modelare şi analiză cantitativă a calităţii cu care serviciile de telecomunicaţii sunt îndeplinite. Este vorba de un instrument amplu de lucru pentru care lucrarea rezervă un număr de şapte capitole pe parcursul cărora se face o prezentare, însoţită de exemplificări în domeniu, a principalelor noţiuni legate de *variabilele şi procesele aleatorii*, staţionaritate şi ergodicitate, *lanţuri Markov şi procese de naştere şi moarte*. Nu sunt lăsate deoparte problemele fundamentale legate de *generarea variabilelor şi funcţiilor aleatorii*, elemente importante în conceperea modelelor de simulare. De asemenea sunt prezentate *metode de aprecieri* relative la proprietăţile statistice ale unui fenomen aleatoriu, prin supunerea acestuia la *diverse teste*, cu ajutorul cărora se pot stabili atât proprietăţile temporare (staţionaritatea, autocorelaţia etc.) cât şi proprietăţi pe ansamblu (media, varianţa, funcţia de repartiţie etc.) ale fenomenului studiat.

Ultima parte, a treia, se ocupă, cum s-a arătat mai sus, cu cele două operaţii fundamentale de studiu al calităţii servirii, şi anume: modelarea şi analiza. Capitolele incluse sunt în număr de trei şi au ca subiecte: *traficul de telecomunicaţii*, pentru care se dau definiţii, se prezintă modele de surse şi se descriu metode de estimare a parametrilor, *sistemele de servire cu aşteptare*, pentru care se precizează care sunt elementele teoretice constitutive şi se dezvoltă un număr apreciabil de modele, şi *reţelele cu şiruri de aşteptare*, privitor la care sunt supuse atenţiei mai multe modele, precum: reţelele deschise fără revenire, cu sau fără pierderi, şi reţelele cu formă produs, de tip Jackson, BCMP sau Kelly, la care se adaugă problematica legată de conceptele de *fiabilitate şi disponibilitate*.

Cartea se adresează, cu precădere, studenţilor din cadrul Facultăţii de Electronică, Telecomunicaţii şi Tehnologia Informaţiei, din Universitatea Politehnica Bucureşti, de la cursurile de licenţă, dar şi de master, în a căror arie curriculară se regăsesc discipline precum: *Ingineria traficului*, *Evaluări de performanţă*, *Proiectarea şi managementul reţelelor radio, trafic şi planificarea frecvenţelor*, *Servicii de telecomunicaţii şi evaluării de performanţă*, *Planificarea serviciilor şi reţelelor*, şi inginerilor din domeniul IT&C.

Aparţinând unei ramuri în care volumul termenilor şi abrevierilor este în continuă expansiune, numărul cuvintelor din vocabularul minimal al oricărei limbi fiind cu mult depăşit, materialul se încheie cu un index necesar unei localizări cât mai facile a noţiunilor specifice subiectului tratat, acolo unde cititorul poate găsi informaţia căutată. Pentru toţi aceşti termeni sunt prezentate, în text şi în index, denumirile în limba engleză, ele fiind, de altfel, folosite aproape în exclusivitate de către toţi specialiştii de telecomunicaţii, indiferent de naţionalitate.

Autorii

30.07.2012

Cuprins

INTRODUCERE **1**

PARTEA I-A: CONCEPTE FUNDAMENTALE **7**

CAPITOLUL 1: FACTORI DETERMINANŢI **9**

1.1 SERVICII DE VOCE, DATE ŞI IMAGINI 9

1.1.1. Raportul semnal-zgomot *10*

1.1.2. Transmiterea vocii *11*

1.1.3. Circuite de date *12*

1.1.4. Video (TV) *12*

1.2 IMPERFECŢIUNILE MEDIILOR DE COMUNICARE 13

1.2.1. Distorsiunea de amplitudine *13*

1.2.2. Distorsiunea de fază *13*

1.2.3. Zgomotul *14*

1.3 CODARE ŞI COMPRESIE 15

1.4 ÎNTÂRZIERI 18

1.4.1. Întârzierea la construirea cadrelor *19*

1.4.2. Întârzierea la corectarea erorilor *20*

1.4.3. Întârzierea la întreţesere *20*

1.4.4. Întârzierea la împachetare *22*

1.4.5. Întârzierea în tamponul de acumulare *23*

1.4.6. Întârzierea în reţea *27*

1.5 PIERDERILE DE PACHETE 30

1.6 EVALUĂRI SUBIECTIVE 34

CAPITOLUL 2: MIJLOACE DE MANAGEMENT AL CALITĂŢII **37**

2.1 MARCAREA PACHETELOR 39

2.2 CLASIFICAREA PACHETELOR 39

2.3 SUPRAVEGHEREA TRAFICULUI 40

2.3.1. Caracteristici de trafic *41*

2.3.2. Contorizarea şi colorarea traficului *42*

2.3.2.1. Marcarea în trei culori pe baza unui sigur debit 42

2.3.2.2. Marcarea în trei culori pe baza a două debite 45

2.4 MANAGEMENTUL COZILOR 46

2.4.1. Mecanismul "tail-drop" *47*

2.4.2. Mecanismul RED *50*

2.4.3. Mecanismul WRED *51*

2.4.4. Mecanismul ECN *52*

2.5 PLANIFICAREA PACHETELOR 54

2.5.1. Procedeul FIFO *55*

2.5.2. Procedeul PQ ... *55*
2.5.3. Procedeul RR ... *56*
2.5.4. Procedeul WRR ... *57*
2.5.5. Procedeul WFQ ... *58*
2.5.6. Procedeul CB-WFQ ... *60*
2.5.7. Procedeul DRR ... *62*
2.6 PROFILAREA TRAFICULUI ... 63
2.6.1. Procedeul Leaky Bucket ... *63*
2.6.2. Procedeul Token Bucket ... *64*
2.7 CONTROLUL ADMITERII ... 66
2.8 PROTOCOALE DE TRANSPORT ÎN TIMP REAL ... 71
2.8.1. Protocolul RTP ... *72*
2.8.2. Protocolul RTCP ... *74*

CAPITOLUL 3: METODE SPECIFICE REŢELELOR IP ... 81
3.1 APROVIZIONAREA ÎN EXCES ... 82
3.2 INTEGRAREA SERVICIILOR ... 86
3.2.1. Clase de servicii ... *87*
3.2.2. Protocolul RSVP ... *89*
3.3 DIFFSERV ... 92
3.3.1. Contractarea serviciilor (SLA/SLS) ... *93*
3.3.2. Arhitectura DiffServ ... *95*
3.3.3. Comportamente caracteristice nodurilor (PHBs) ... *99*
3.3.3.1. Înaintare expeditivă ... 100
3.3.3.2. Înaintare asigurată ... 102
3.4 INTSERV PESTE DIFFSERV ... 103

PARTEA II-A: PROBABILITĂŢI ŞI PROCESE ALEATORII ... 105

CAPITOLUL 4: EXPERIMENT ŞI MODELARE ... 107
4.1 EXPERIMENTE ALEATORII ... 109
4.2 AXIOMELE PROBABILITĂŢII ... 111
4.3 PROBABILITĂŢI CONDIŢIONATE ... 114
4.4 INDEPENDENŢA EVENIMENTELOR ... 118
4.5 EXPERIMENTE SECVENŢIALE ... 119
4.6 APLICAŢII SUPLIMENTARE ... 124

CAPITOLUL 5: VARIABILE ALEATORII - definiţii şi concepte de bază ... 131
5.1 FUNCŢIA DE DISTRIBUŢIE ... 132
5.2 FUNCŢIA DE DENSITATE DE PROBABILITATE ... 133
5.3 FUNCŢII DE O VARIABILĂ ALEATORIE ... 139
5.4 MĂRIMI CARACTERISTICE ALE UNEI VARIABILE ALEATORII ... 140
5.5 TRANSFORMATELE UNEI VARIABILE ALEATORII ... 142
5.6 DISTRIBUŢII UZUALE DE VARIABILE ALEATORII ... 145
5.6.1. Distribuţia Bernoulli ... *146*
5.6.2. Distribuţia binomială ... *146*
5.6.3. Distribuţia geometrică ... *146*
5.6.4. Distribuţia binomială negativă ... *147*
5.6.5. Distribuţia Poisson ... *148*
5.6.6. Distribuţia uniformă discretă ... *148*
5.6.7. Distribuţia uniformă continuă ... *149*

5.6.8. Distribuţia exponenţială *149*
5.6.9. Distribuţia Pareto *150*
5.6.10. Distribuţia gaussiană (normală) *150*
5.6.11. Distribuţia Gama *153*
5.6.12. Distribuţia m-Erlang *154*
5.6.13. Distribuţia χ^2 ..... *154*
5.6.14. Distribuţia Student, t *156*
5.6.15. Distribuţia Fisher - Snedecor ($\mathcal{F}$ *)* *156*
5.6.16. Distribuţia Laplace *157*
5.6.17. Distribuţia Rayleigh *157*
5.6.18. Distribuţia Cauchy *157*
5.7 METODE DE APROXIMARE 158
5.7.1. Inegalitatea Markov *158*
5.7.2. Inegalitatea Cebâşev *158*
5.7.3. Inegalitatea Chernoff *159*
5.7.4. Marginea reuniunii *160*
5.7.5. Calculul aproximativ al valorilor aşteptate *161*
5.7.5.1. Dezvoltarea în serie Taylor trunchiată 161
5.7.5.2. Aproximarea prin cuadratură 164
5.8 APLICAŢII SUPLIMENTARE 165

CAPITOLUL 6: VARIABILE ALEATORII VECTORIALE 171
6.1 VARIABILE ALEATORII BIDIMENSIONALE 173
6.1.1. Independenţa componentelor *180*
6.1.2. Probabilităţi condiţionate şi mărimi caracteristice condiţionate *182*
6.2 VARIABILE ALEATORII n -DIMENSIONALE ($n > 2$) 186
6.3 FUNCŢII DE MAI MULTE VARIABILE ALEATORII 188
6.3.1. Mărimi caracteristice ale funcţiilor de două variabile aleatorii *189*
6.3.2. Sume de variabile aleatorii *191*
6.3.3. Media pe eşantion *194*
6.3.4. Legea slabă a numerelor mari *194*
6.3.5. Legea tare a numerelor mari *196*
6.3.6. Teorema limită centrală *197*
6.4 INTERVALE DE ÎNCREDERE 202
6.4.1. Noţiuni de bază *202*
6.4.2. Determinarea intervalelor de încredere *202*
6.5 METODA FAZELOR 206
6.5.1. Structură tip serie *207*
6.5.2. Structură tip paralel *208*
6.5.3. Structură tip serie-paralel *209*
6.5.4. Structură tip Cox *210*
6.6 FUNCŢII VECTORIALE DE VARIABILE ALEATORII 211
6.6.1. Transformări liniare *212*
6.6.2. Transformări neliniare *214*
6.7 APLICAŢII SUPLIMENTARE 217

CAPITOLUL 7: GENERAREA NUMERELOR ALEATORII 223
7.1 METODE DE GENERARE A NUMERELOR ALEATORII DISTRIBUITE UNIFORM 223
7.2 METODE DE GENERARE A NUMERELOR ALEATORII DISTRIBUITE ARBITRAR 227
7.2.1. Metoda transformării analitice *227*
7.2.2. Metoda respingerii *232*

7.2.3. Generarea funcţiilor de variabile aleatorii ... *235*
7.3 APLICAŢII SUPLIMENTARE ... 238

CAPITOLUL 8: TESTE STATISTICE ... 241
8.1 TESTUL DE STAŢIONARITATE ... 241
8.2 TESTUL DE AUTOCORELAŢIE ... 241
8.3 HISTOGRAMA ... 243
8.4 TESTUL χ^2 ... 245
8.5 TESTUL KOLMOGOROV-SMIRNOV ... 248
8.6 APLICAŢII SUPLIMENTARE ... 254

CAPITOLUL 9: PROCESE ALEATORII ... 257
9.1 NOŢIUNI INTRODUCTIVE ... 257
9.2 SPECIFICAREA UNUI PROCES ALEATORIU ... 259
9.3 STAŢIONARITATEA PROCESELOR ALEATORII ... 261
9.4 MEDII TEMPORALE ŞI CONDIŢII DE ERGODICITATE ... 268
9.5 PROCESE AUTOSIMILARE ... 273
9.6 PROCESE ALEATORII TIP SUMĂ ... 278
9.7 PROCESE AR (p) ... 284
9.8 PROCESE DE MEDIERE PRIN ALUNECARE MA (q) ... 286
9.9 PROCESE ARMA (p,q) ... 286
9.10 PROCESE ARIMA (p,d,q) ... 287
9.11 PROCESE f ARIMA (p,d,q) ... 288
9.12 PROCESE GAUSSIENE ... 289
9.13 APLICAŢII SUPLIMENTARE ... 291

CAPITOLUL 10: LANŢURI MARKOV ... 295
10.1 LANŢURI MARKOV DISCRETE ... 297
10.2 LANŢURI MARKOV CONTINUE ... 305
10.3 METODA VALORILOR ŞI VECTORILOR PROPRII ... 312
10.3.1. Valori şi vectori proprii ... 312
10.3.2. Calculul probabilităţilor de stare în lanţuri Markov discrete ... 313
10.3.3. Calculul probabilităţilor de stare în lanţuri Markov continue ... 315
10.4 PROCESE DE NAŞTERE ŞI MOARTE ... 318
10.5 CVASI PROCESE DE NAŞTERE ŞI MOARTE ... 321
10.5.1. Metoda matriceal-geometrică ... 321
10.5.2. Aplicarea procedurii MGM în cazul proceselor QBD ... 322
10.6 APLICAŢII SUPLIMENTARE ... 330

PARTEA III-A: MODELARE ŞI ANALIZĂ ... 333

CAPITOLUL 11: TRAFICUL DE TELECOMUNICAŢII ... 335
11.1 DEFINIŢII, UNITĂŢI ŞI METODE DE MĂSURARE ... 335
11.2 SURSE DE TRAFIC ... 343
11.2.1. Clasificare ... 344
11.2.2. Modelarea surselor de trafic ... 347
11.2.2.1. Surse de trafic ON/OFF ... 348
11.2.2.2. Surse de trafic (continuu) variabil ... 351
11.3 ESTIMAREA PARAMETRILOR ... 355
11.3.1. Verosimilitatea maximă a datelor ... 356

11.3.2. Algoritmul EM ... *360*
11.3.3. Studii de caz ... *360*
11.3.3.1. Model de amestec ... 360
11.3.3.2. Lanţuri Markov ascunse ... 364
11.3.3.3. Distribuţii tip fază ... 370
11.3.3.4. Procese Markov de sosiri ... 381
11.3.3.5. Procese Markov de sosiri ponderate ... 387
11.4 MODELAREA MOBILITĂŢII ... 390
11.5 APLICAŢII SUPLIMENTARE ... 397
CAPITOLUL 12: SISTEME DE SERVIRE CU AŞTEPTARE ... 401
12.1 ELEMENTELE DE MODELARE ... 403
12.2 FORMULA LUI LITTLE ... 404
12.3 CLASIFICAREA KENDALL ... 407
12.4 MODELUL M/M/1 ... 409
12.5 MODELUL M#/M/1 CU CONTROL DE FLUX ... 413
12.6 MODELUL M/M/1 / ∞ / *r* / FIFO ... 415
12.7 MODELUL M/M/ *s* ... 418
12.8 MODELUL M/M/ *s* / ∞ / *s* CU PIERDERI ... 421
12.9 MODELUL REVĂRSĂRII TRAFICULUI ... 423
12.9.1. Grup cu revărsare primară ... *423*
12.9.2. Fascicul comun de revărsare ... *429*
12.10 MODELUL M/M/ *s* / ∞ / *r* / FIFO ... 432
12.11 MODELUL M/M/1/ *c* ... 435
12.12 MODELUL M/G/1 ... 441
12.12.1. Serviri cu întrerupere ... *447*
12.12.2. Serviri cu prioritate ... *449*
12.12.2.1 Sisteme fără întreruperea servirii curente ... 450
12.12.2.2. Sisteme cu întreruperea servirii curente ... 452
12.12.3. Modelul Geom/G/1 ... *453*
12.13 MODELUL GI/M/1 ... 454
12.14 MODELUL GEOM[*A*]/D/1 ... 455
12.14.1. Ecuaţia de evoluţie şi valorile medii ... *456*
12.14.2. Funcţia generatoare de probabilităţi ... *458*
12.15 INTEGRAREA SERVICIILOR ... 459
12.15.1. Integrarea utilizând disciplina FIFO ... *461*
12.15.2. Integrarea cu prioritate absolută ($s=1$) ... *464*
12.15.3. Integrarea cu rezervare şi graniţă mobilă ... *466*
12.16 MODELE FLUIDE ... 471
12.16.1. Modelare matematică ... *472*
12.16.1.1. Modelul fluid cu acumulare ... 472
12.16.1.2. Modelul fluid fără acumulare ... 484
12.16.2. Studii de caz ... *486*
12.17 APLICAŢII SUPLIMENTARE ... 488
CAPITOLUL 13: REŢELE DE ŞIRURI CU AŞTEPTARE ... 503
13.1 INTRODUCERE ... 503
13.2 DEFINIŢII ŞI NOTAŢII ... 506
13.3 INDICATORI DE PERFORMANŢĂ ... 509
13.4 TEOREMA LUI BURKE ... 511
13.5 REŢELE DESCHISE FĂRĂ REVENIRE ... 513
13.5.1. Rutare deterministă ... *515*

13.5.2. Rutare probabilistică 519
13.6 REŢELE FĂRĂ REVENIRE, CU PIERDERI DISTRIBUITE 520
13.7 REŢELE CU FORMĂ PRODUS 523
13.7.1. Reţele Jackson 524
13.7.2. Reţele Gordon-Newell 530
13.7.2.1. Teorema lui Gordon şi Newell 531
13.7.2.2. Algoritmul Buzen 535
13.7.2.3. Metoda valorii medii 536
13.7.3. Reţele BCMP 537
13.7.3.1. Clase şi lanţuri de clienţi 538
13.7.3.2. Forme produs în reţele BCMP 539
13.7.3.3. Metode aproximative de rezolvare a reţelelor BCMP 544
13.7.4. Reţele Kelly 549
13.7.4.1. Reţele Kelly deschise 549
13.7.4.2. Reţele Kelly închise 552
13.8 DISPONIBILITATEA REŢELELOR DE TELECOMUNICAŢII 553
13.8.1. Fiabilitate şi disponibilitate 553
a) Proiectarea redundantă 555
b) Reconfigurarea 556
c) Detectarea, localizarea şi semnalarea defecţiunii 556
13.8.2. Evaluarea disponibilităţii unui sistem de telecomunicaţii 557
13.8.2.1. Sistem elementar 557
13.8.2.2. Sisteme compuse 559
13.8.3. Evaluarea disponibilităţii reţelelor 565
13.8.3.1. Metoda grafo-analitică de evaluare 566
13.8.3.2. Evaluarea disponibilităţii prin modelare de tip Markov 570
13.8.4. Aplicaţii suplimentare 573
13.9 CONCLUZII GENERALE 575
13.10 APLICAŢII SUPLIMENTARE 577

INDEX 583

BIBLIOGRAFIE 589

Introducere

Informaţia, într-un anumit context, este tot ceea ce are sens pentru cine o recepţionează, fiind prezentă în diferite forme precum: *text, desen, mesaj vocal, imagine fixă sau în mişcare şi prezentare multimedia*. În cazul reţelelor de comunicaţii şi ale sistemelor de calcul, informaţia este **generată**, **colectată**, **prelucrată**, **înmagazinată**, **transferată** şi **utilizată** în diverse forme, precum cele menţionate anterior, toate aceste activităţi apelând la un ansamblu vast şi dinamic de mijloace tehnice de comunicare şi calcul. Referirea la acest ansamblu se face prin termenul general de **tehnologia informaţiei** sau, în exprimare completă, ***tehnologia informaţiei şi comunicării***, pentru a evidenţia amploarea celor două subcategorii de mijloace, mai mult sau mai puţin şi, în fapt, tot mai puţin independente!

Informaţia este produsă de oameni, aparate sau orice poate fi perceput, la modul abstract, drept proces şi **este recepţionată** de alte procese care la rândul lor, în urma unor prelucrări caracteristice, generează informaţie nouă, lanţul continuând la nesfârşit. Realizarea acestui schimb informaţional implică existenţa unui **mediu de comunicare** (interconectare), prin care informaţia este transportată (transferată). Termenul **transport** este provenit din două cuvinte ale limbii latine: *trans*, tradus în româneşte ca "prin", şi *portare*, tradus cu "a purta") şi reprezintă transferul informaţiei de la generator (sursă, expeditor) la receptor (destinatar) prin intermediul unui element purtător (*carrier*). Purtătorul poate fi reprezentat, de exemplu, de undele acustice în cazul unui dialog verbal direct între două sau mai multe persoane, sau de către un furnizor oarecare al unui *serviciu de livrare*, de exemplu de genul **poştă**, atunci când informaţia este înregistrată pe hârtie şi transportată în plic, sau de genul **furnizor de servicii Internet ISP** (***I**nternet **S**ervice **P**rovider*), atunci când informaţia este scrisă şi apoi trimisă prin intermediul unei ferestre de compunere (*composer*) a unei aplicaţii de poştă electronică, de exemplu **SMNP** (***S**imple **M**ail **N**etwork **P**rotocol*).

Pe lângă informaţie, schimburile şi transferurile au în vedere şi alte "obiecte" care, în general, nu suportă transportul prin traducerea lor într-o formă abstractă, descrisă prin intermediul unui set de simboluri. Astfel de obiecte sunt bunurile materiale şi oamenii pentru care civilizaţia umană a inventat şi dezvoltat diverse reţele de comunicaţii, precum sunt reţelele de transport feroviar, auto, aerian sau maritim, teleportarea fiind în continuare de domeniul ficţiunii ştiinţifice.

Transportul, în general, şi al informaţiei, în particular, este descris de o serie de caracteristici, printre care: **întârzieri**, **pierderi** şi **blocări**/**respingeri**, care depind atât de însuşirile mediului de comunicare, cum ar fi: **capacitatea** sau **disponibilitatea**, cât şi de nevoile de schimb de informaţie, de exemplu: **frecvenţa** sau **cantitatea**. Toate aceste însuşiri se regăsesc în conceptul, în fapt larg utilizat şi în vorbirea curentă, denumit **trafic**, care constituie obiectul de studiu al unei ample teorii intitulată: **teoria aşteptării** (*Queueing Theory*).

O **reţea de comunicaţie** este un ansamblu de mijloace prin care se poate comunica între mai multe locaţii distincte. Comunicarea presupune existenţa, atât a unor ***căi de comunicare*** (de *interconectare*), cât şi a unor ***mijloace de circulaţie*** (transport) specifice. Astfel, de exemplu, în cazul reţelelor de comunicaţii feroviare avem căi ferate şi trenuri, în cazul reţelelor de transport auto avem drumuri, străzi, şosele şi autostrăzi pe care circulă autoturisme, autobuze, camioane etc., iar în cazul reţelelor de transport aeriene avem aeroporturi şi culoare de zbor pentru avioane sau elicoptere.

În ceea ce priveşte informaţia, se poate recurge la reţeaua de comunicaţii poştale, care pentru transport utilizează reţelele menţionate anterior, fiind astfel simultan furnizor şi beneficiar (utilizator) de servicii, sau se pot folosi reţelele de telecomunicaţii şi/sau reţelele de calculatoare, unde căile de comunicare sunt mediile de propagare, iar mijloacele de transport (purtător) sunt undele electromagnetice.

Reţelele de comunicaţii, de genul reţelelor de telecomunicaţii şi/sau reţelelor de calculatoare, sunt aranjamente de componente materiale (**hardware**) şi logice (**software**) care oferă o serie de servicii de comunicare permiţând schimbul de informaţii între utilizatorii săi. Utilizatorii sunt, de obicei, oameni, dar şi echipamente, dispozitive sau programe de calculator.

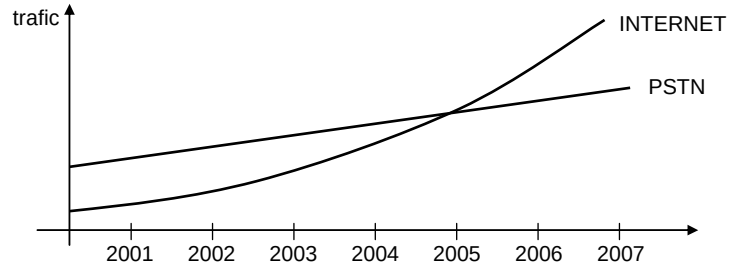


Figura 1: Evoluţia traficului în PSTN şi INTERNET

Una dintre cele mai familiare şi răspândite (încă!) reţele de comunicaţii este reţeaua publică de telefonie fixă **PSTN** (***P**ublic **S**witched **T**elephone **N**etwork*). Această reţea a fost destinată iniţial pentru a oferi serviciul de telefonie **POTS** (***P**lain **O**ld **T**elephone **S**ervice*), dar, pe parcurs, a fost adaptată pentru noi servicii precum: identificarea liniei chemătoare (*calling line identification*) sau accesul la Internet. Un alt exemplu de reţea de comunicaţie îl constituie reţelele particulare de calculatoare (*office computer network*) care au apărut ca urmare a

răspândirii calculatoarelor personale şi a staţiilor de lucru şi nevoii de conectare a acestora în scopul partajării informaţiei (programe şi date) şi a echipamentelor periferice (imprimante, scanere etc.). Cel din urmă exemplu îl constituie **Internetul** care a depăşit reţeaua PSTN în popularitate şi utilizare (vezi figura 1). Internetul este de fapt o reţea de reţele de calculatoare, care acoperă aproape întreaga suprafaţă a globului şi care permite sutelor de milioane de utilizatori să schimbe între ei informaţii de orice natură, anume de tip *text, audio sau video*, precum şi combinaţii ale acestora, în servicii *multimedia*.

Proiectarea reţelelor de comunicaţii are ca obiective principale stabilirea configuraţiei, fixarea capacităţii legăturilor de comunicaţie şi identificarea unui mod optim prin care acestea sunt partajate între fluxurile informaţionale active la un moment dat. Toate aceste sarcini trebuie raportate, deopotrivă, la operatorii de reţea şi la utilizatorii de servicii. Punctul de vedere al operatorilor se referă, cu precădere, la asigurarea unei utilizări cât mai eficiente a resurselor şi la satisfacerea unui număr cât mai mare de transferuri informaţionale, scopul final fiind obţinerea unor câştiguri maxime cu cheltuieli minime. În schimb, pretenţiile utilizatorilor vizează, în principal, modul în care serviciile sunt satisfăcute, situaţii precum aşteptarea unei căi libere până la destinaţie, pierderea sau alterarea informaţiei şi întârzierea transferurilor fiind de dorit a fi cât mai rare. Pentru toate aceste aspecte sunt definite următoarele două categorii de indicatori, ce caracterizează serviciile de comunicaţii:

- **gradul de servire, GoS** (***G****rade* ***o****f* ***S****ervice*), care se referă la posibilităţile de a beneficia de un anumit serviciu;
- **calitatea servirii, QoS** (***Q****uality* ***o****f* ***S****ervice*), care precizează posibilităţile pe care le oferă un anumit serviciu pe durata utilizării acestuia.

În cazul reţelelor cu comutaţie de circuite, destinate, cu precădere, traficului telefonic, obiectivele majore care au guvernat implementarea lor au fost menţinerea în limite rezonabile a **probabilităţii de blocare/respingere a cererilor de serviciu** şi a **disponibilităţii serviciilor** în cauză (ca **indicatori ai gradului de servire**) şi optimizarea planului de transmisie a semnalului vocal, care a vizat o serie de **indicatorii de calitate a servirii**, legaţi de **pierderi**, **zgomot**, **ecou**.

Spre deosebire de reţelele cu comutaţie de circuite, reţelele cu comutaţie de pachete au avut în vedere, iniţial, transmisia de date care, fiind un serviciu fără restricţii temporale, în comparaţie cu serviciile în timp real, de tip voce sau video, s-a putut implementa adoptând principiul **înmagazinării şi înaintării** (***stored and forward***) cu posibilitatea retransmisiei în caz de eroare. Mai mult, în cazul reţelelor IP, unde unul dintre principiile fundamentale, de implementare este cel al concentrării inteligenţei la periferie, luarea în considerare doar a **terminale**lor **inteligente** (*host computers*) a permis ca, prin degrevarea de o serie de sarcini, acestea să fie realizate, iniţial, cât mai simplu (spre deosebire de reţelele de telefonie care concentrau inteligenţa în interiorul lor, de unde şi conceptul de ***Intelligent Network***). revenindu-le, în principal, doar sarcina de a înainta pachetele de la un nod la altul, fără a face vreo diferenţiere între ele. În

aceste condiţii, transmiterea datelor, ca unic serviciu oferit, este catalogat, din punct de vedere al calităţii, ca fiind de tip ***"cât mai bine posibil"*** (***best effort***).

În prezent, etapa reţelelor specializate este pe cale de a se încheia, tendinţa actuală fiind convergenţa, implementarea unor reţele unice, care să îmbine şi să se adapteze în mod eficient metodele folosite în cadrul celor două tipuri de reţele, precum ingineria traficului din reţelele cu comutaţie de circuite şi multiplexarea statistică, proprie reţelelor cu comutaţie de pachete, şi să adopte noi tehnologii, astfel încât să fie capabilă să asigure servicii de orice gen, de date, audio şi video. Desigur, o astfel de reţea, bazată pe comutaţia de pachete, nu se mai poate rezuma doar la mecanismul "*best effort*", insuficient pentru a satisface diverse cerinţe, deseori conflictuale, caracteristice tipurilor diferite de informaţii transferate printr-o astfel de reţea (vezi tabelul 1). Dimpotrivă, asigurarea calităţii, în condiţiile unei diversităţi de servicii, presupune ca reţeaua în cauză să dispună de mecanisme (tehnologii) capabile să diferenţieze traficul în clase şi să prelucreze în mod specific membrii fiecărei clase, dar luând în considerare şi ansamblul.

Tabelul 1: Exemple de servicii şi niveluri ale cerinţelor asociate

Application	Reliability	Delay	Jitter	Bandwidth
E-mail	High	Low	Low	Low
File Transfer	High	Low	Low	Medium
Web access	High	Medium	Low	Medium
Remote login	High	Medium	Medium	Low
Audio on Demand	Low	Low	High	Medium
Video on Demand	Low	Low	High	High
Telephony	Low	High	High	Low
Videoconferencing	Low	High	High	High

Definirea mijloacelor şi metodelor destinate asigurării calităţii servirii (*QoS mechanisms*), obiectiv al primei părţi din prezenta lucrare, implică o serie de activităţi de modelare analitică sau experimentală (prin simulare, de exemplu), şi analiză, care, la rândul lor, apelează la diverse discipline matematice, precum: teoria probabilităţilor şi statistică, în vederea stabilirii modalităţilor de acţionare, luând în considerare factorii care au influenţă asupra calităţii servirii. Dispunând de un bagaj teoretic amplu, cu caracter general, fiecare din activităţile şi instrumentele amintite constituie un subiect separat, a cărui expunere revine în sarcina următoarelor două părţi ale acestei lucrării. Este vorba de partea a II-a, care are ca obiectiv prezentarea instrumentelor analitice fundamentale folosite în descrierea fenomenelor aleatorii, şi de partea a III-a, care se ocupă cu aspectele specifice serviciilor de telecomunicaţii, privitoare la activităţile de modelare şi analiză a calităţii acestora..

Indicatorii utilizaţi pentru exprimarea calităţii servirii în reţelele cu comutaţie de pachete se împart în două categorii: **categoria indicatorilor asociaţi serviciilor cu sau fără conexiune** şi **categoria indicatorilor proprii serviciilor cu conexiune**.

În **prima categorie** se află indicatorii legaţi de transmiterea unei singure unităţi informaţionale (pachet, celulă, cadru), precum:

- întârzierea măsurată prin medie, deviaţie standard etc.,
- probabilitatea coruperii,
- probabilitatea pierderii,
- probabilitatea duplicării,
- probabilitatea livrării greşite,
- protecţia faţă de accesul neautorizat,
- prioritatea în livrare.

dar şi parametri care iau în considerare fluxurile informaţionale în ansamblu, ca de exemplu:

- eficienţa transferului, măsurată prin medie, deviaţie standard etc.,
- probabilitatea livrării în dezordine.

Din **a doua categorie** fac parte parametrii precum:

- întârzierea stabilirii unei conexiuni,
- probabilitatea eşecului în stabilirea unei conexiuni,
- întârzierea eliberării unei conexiuni,
- elasticitatea conexiunii.

Evaluarea indicatorilor de performanţă, de genul celor mai sus menţionaţi, se poate face, conform celor menţionate mai sus, prin **modelare matematică** sau **analiză statistică** a rezultatelor experimentale. În primul caz, beneficiem de formule de calcul, valabile în anumite ipoteze, ca de exemplu, considerarea procesului de sosire a pachetelor într-o interfaţă de ieşire a unui anumit ruter ca fiind de tip Poisson. În al doilea caz, uzăm de un aparat matematic diversificat, care ne oferă o serie de **statistici** (medie, deviaţie standard, autocovarianţă, interval de încredere etc.) corespunzătoare eşantionului de rezultate obţinute prin măsurători.

Măsurătorile se pot desfăşura într-un cadru real de funcţionare a sistemului sau prototipului supus analizei, ori într-un context virtual, folosind simularea. Principial, **simularea** constă în imitarea, prin intermediul, de exemplu, a unui program pe calculator, a realităţii, atât în privinţa modului de funcţionare a sistemului, cât şi a sarcinii la care este supus.

Pe lângă mijloacele mai sus prezentate, evaluările de performanţă pot apela şi la o soluţie hibridă, intitulată **emulare**. În acest caz, experimentul vizează un ansamblu compus dintr-o componentă reală şi una virtuală. De exemplu, componenta reală poate fi reprezentată de două sau mai multe entităţi concrete, localizate în aceeaşi gazdă (*host*) sau în gazde distante, care corespondează pe bază unui protocol supus analizei, iar componenta virtuală o poate constitui mijlocul de comunicare utilizat, implementat software şi/sau hardware, ale cărui imperfecţiuni, în speţă întârzieri şi pierderi, se manifestă în conformitate cu un anumit model matematic.

Indicatorii de calitate/performanţă enumeraţi mai sus sunt specifici nivelurilor legătură de date şi reţea, din cadrul modelului OSI. Pe lângă aceştia, calitatea serviciilor de telecomunicaţii mai este determinată de două mărimi ce caracterizează nivelul fizic, şi anume: raportul semnal zgomot şi frecvenţa biţilor eronaţi. Despre aceste mărimi se fac precizări în cele ce urmează, în contextul claselor specifice de servicii, amintindu-se, de asemeni, influenţe nedorite, induse de componentele reţelei asupra lor. Nu se va insista asupra detaliilor, cei interesaţi putând consulta lucrări pe profil, de genul [Far], [IoNic], [Ioan98].

capitolul 4

Experiment şi modelare

În tentativa sa de a-şi explica fenomenele cu care intră în contact, fiinţa umană, mărginită prin natura sa, încearcă să simplifice mental complexitatea acestora fără a prejudicia esenţialul. Astfel, se obţin nişte **modele (paradigme)** prin care se urmăreşte o cât mai bună corelare a informaţiilor deţinute la un moment dat.

Modelele de care dispune o societate stau la baza producerii şi dezvoltării diverselor bunuri materiale şi spirituale ale acesteia. De cele mai multe ori, realizarea sau dezvoltarea acestor bunuri este precedată de o etapă de căutare a soluţiei optime. În general, stabilirea soluţiei optime se face comparând evaluările cantitative ale diverşilor indicatori precum: costul (investiţiile şi cheltuielile de amortizare), disponibilitatea (probabilitatea funcţionării la un moment dat), flexibilitatea (capacitatea de adaptare la evoluţia factorilor externi şi interni), performanţa (gradul de satisfacere a obiectivelor propuse), ce caracterizează diversele variante luate în calcul. Evaluările cantitative se pot obţine prin *observarea* unor realizări experimentale sau construind şi analizând **modele** pentru variantele considerate. Deseori, a doua metodă este de preferat măsurătorilor experimentale ce pot implica un cost excesiv (nejustificat).

Un **model** este o reprezentare aproximativă care descrie şi explică, cu ajutorul unui set de reguli simple şi intuitive, principalele aspecte ce caracterizează un anumit fenomen.

Un model matematic este un ansamblu de relaţii matematice ce leagă principalii parametrii şi variabile ce caracterizează fenomenul studiat. Soluţia acestor relaţii reprezintă previziunea măsurătorilor care ar fi obţinute dacă experimentul ar avea loc. Condiţia necesară (nu şi suficientă) de dezvoltare a unui model matematic este ca fenomenul studiat să prezinte *însuşiri măsurabile.*

Atunci când nu se poate dezvolta un model matematic, se poate recurge la *modele de simulare computerizate*. Un **model de simulare computerizat** este un program a cărui execuţie "imită", la un nivel de detaliu limitat de performanţele maşinilor utilizate, evoluţia (dinamica) fenomenului studiat. Cu ajutorul unor

instrucţiuni care înregistrează valorile variabilelor caracteristice fenomenului, execuţia modelului de simulare computerizat generează un set de date primare. Aceste date sunt supuse unei analize statistice ce oferă previziuni relative la fenomenul studiat.

În general, simulările sunt capabile să descrie realitatea mult mai precis decât modelele matematice. În schimb, simulările sunt mult mai puţin flexibile şi timpii de execuţie sunt de obicei mai mari decât în cazul modelelor matematice.

Corectitudinea şi precizia cu care un model aproximează o situaţie dată se controlează printr-un **proces de modelare** (investigaţie, descriere, specificare, validare şi dezvoltare). Acest proces constă într-o serie de experimente şi modificări ale modelului, urmărind algoritmul prezentat în figura 4.4. Un experiment definit cu scopul verificării ipotezelor investighează un anumit aspect al fenomenului studiat. Experimentul se desfăşoară în anumite condiţii şi efectele sale sunt observate şi măsurate. Modelul corespunzător oferă previziuni legate de rezultatele experimentului. Dacă apar discrepanţe semnificative, modelul şi, eventual, ipotezele se revizuiesc (modifică).

Procesul de modelare continuă până când evoluţia tuturor aspectelor importante ale fenomenului poate fi prezisă de model cu o precizie dorită. Alegerea aspectelor importante depinde de obiectivele urmărite de investigator. Din acest motiv, pentru acelaşi fenomen se pot construi modele distincte, care oferă previziuni pentru mărimi diferite şi sunt valabile în anumite condiţii de desfăşurare a fenomenului.

A treia ramură a procesului de modelare, **simularea**, este necesară atunci când construirea unui model matematic este dificilă şi când obiectul de studiu (produsul, sistemul etc.) se află doar în faza de proiectare (nu există încă!). Astfel, simularea oferă nişte estimări care pot fi interpretate ca atare sau

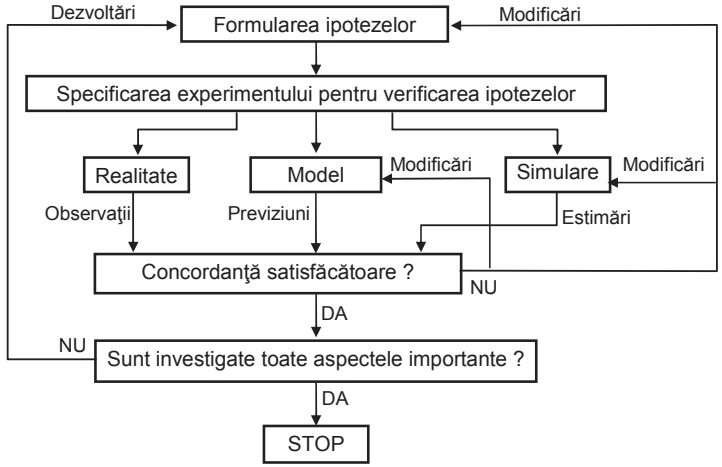


Figura 4.1: Procesul de modelare

pot controla procesul de elaborare a unui model matematic al unei implementări ipotetice, fiind, până în momentul materializării acesteia, unicul său mijloc de validare.

Un **model matematic** este **determinist** atunci când condiţiile de efectuare a experimentului vizat determină exact mărimile (variabilele) urmărite. Teoria circuitelor electrice este un astfel de model. Ea presupune că interacţiunea dintre componentele *idealizate* (discrete, cu caracteristici tensiune/curent ideale) ale unui circuit este complet descrisă de legile lui Kirchhoff (ale tensiunilor şi curenţilor). În practică, există variaţii între mărimile aşteptate (prezise) şi cele observate (măsurate). Aceste diferenţe se datorează erorilor de măsurare şi factorilor neluaţi în considerare de model (variaţia temperaturii, îmbătrânirea componentelor etc.). Totuşi, atât timp cât aceste deviaţii sunt nesemnificative, acest tip de model este cel mai adecvat (prin simplitatea şi intuitivitatea sa) în analiza şi sinteza circuitelor.

4.1 Experimente aleatorii

Un experiment (fenomen) a cărui evoluţie diferă semnificativ atunci când este repetat în aceleaşi condiţii se numeşte **experiment aleatoriu.** Specificarea unui experiment aleator constă în stabilirea procedurii de desfăşurare a sa, precum şi a setului de măsurători şi observaţii ce-l însoţesc. Un experiment aleator poate fi modelat dacă evoluţia sa se caracterizează prin **regularitate statistică** (de exemplu, mediile valorilor obţinute în secvenţe lungi de repetări ale unui experiment au, aproximativ, aceeaşi valoare). Rezultatul indivizibil (ce nu mai poate fi descompus) obţinut în urma desfăşurării unui experiment se numeşte **realizare** sau **eşantion punct** (notat ζ). **Spaţiul realizărilor** sau **spaţiul de eşantionare** (notat S) al unui experiment aleator este mulţimea tuturor realizărilor posibile. Din considerente teoretice este util ca spaţiul realizărilor să includă şi realizările imposibile. În statistică, acest concept poartă denumirea de **populaţie**. Mulţimea realizărilor unui experiment care este finită sau infinit numărabilă (între realizări şi mulţimea numerelor naturale se poate stabili o corespondenţă biunivocă) poartă denumirea de **spaţiu de eşantionare discret**. Mulţimea realizărilor unui experiment care este infinit nenumărabilă se numeşte **spaţiu de eşantionare continuu**.

Atunci când o realizare presupune mai multe observaţii şi/sau măsurători, spaţiul de eşantionare este **multidimensional**. Un astfel de spaţiu poate fi exprimat printr-un produs cartezian de diferite mulţimi.

Se numeşte **eveniment** (notat A) o submulţime a spaţiului de eşantionare. Dacă $A=\mathrm{S}$ vorbim despre un **eveniment sigur.** Dacă $A=\varnothing$ vorbim despre un **eveniment imposibil** sau **eveniment nul.** Într-un spaţiu de eşantionare discret, un eveniment ce constă dintr-o singură realizare se numeşte **eveniment elementar.**

Fiind mulţimi, evenimentele se pot combina cu ajutorul operaţiilor cu mulţimi (reuniunea, intersecţia, complementaritatea etc.) sau se pot compara cu ajutorul relaţiilor dintre mulţimi (incluziune, egalitate).

Aplicaţia 4.1.1

Fie următoarele experimente:

- E_1: la recepţie se detectează trei simboluri; se notează secvenţa binară corespunzătoare;
- E_2: la recepţie se detectează trei simboluri; se numără simbolurile de 1 conţinute (experimentele E_1 şi E_2 se aseamănă prin procedură, dar diferă prin măsurători);
- E_3: un nod de comunicaţie deserveşte N terminale; se notează numărul de terminale active la un moment dat;
- E_4: un bloc de informaţie este repetat la transmisie până când se primeşte un mesaj de confirmare; se numără repetările;
- E_5: la un centru de distribuire sosesc mesaje; se măsoară durata între două sosiri succesive;
- E_6: o componentă este introdusă într-un montaj ce se pune în funcţiune; se măsoară timpul de funcţionare al componentei;
- E_7: la recepţie se urmăreşte semnalul din linie; se măsoară şi se notează tensiunea la momentul t_1;
- E_8: la recepţie se urmăreşte semnalul din linie; se măsoară şi se notează tensiunile la momentele t_1 şi t_2.

i) Să se determine spaţiile de eşantionare corespunzătoare;

ii) Considerând experimentul E_7 se definesc următoarele evenimente:

A - amplitudinea tensiunii este mai mare de 10V,
B - tensiunea este mai mică de –5V,
C - tensiunea este pozitivă.

iii) Să se specifice în limbaj matematic evenimentele:

A, B, C, $A\cap B$, $A\cup B$, C^C, $(A\cup B)\cap C$, $A\cap B\cap C$, $(A\cup B)^C$,

unde $(*)^C$ reprezintă complementul evenimentului $(*)$.

Rezolvare:

i) $S_1=\{000,001,010,100,011,110,101,111\}$; $S_2=\{0,1,2,3\}$;
$S_3=\{0,1,\cdots,N\}$; $S_4=\{0,1,2,3,\cdots\}$; $S_5=S_6=\{t:t\ge 0\}=[0,+\infty)$;
$S_7=\{u:-\infty<u<+\infty\}=(-\infty,+\infty)$;
$S_8=\{(u_1,u_2):-\infty<u_1<+\infty,-\infty<u_2<+\infty\}=\mathrm{R}\times\mathrm{R}$

ii) $A=\{u:u>10\}=(-\infty,-10)\cup(10,+\infty)$;
$B=\{u:u<-5\}=(-\infty,-5)$;
$C=\{u:u>0\}=(0,+\infty)$;

$A \cup B = \{u : u < -5 \text{ sau } u > 10\} = (-\infty, -5) \cup (10, +\infty)$

$A \cap B = \{u : u < -10\} = (-\infty, -10)$; $C^C = \{u : u \le 0\} = (-\infty, 0]$;

$(A \cup B) \cap C = \{u : u > 10\} = (10, +\infty)$; $A \cap B \cap C = \varnothing$;

$(A \cup B)^C = \{u : -5 \le u \le 10\} = [-5, 10]$

* * *

Frecvenţa relativă (notată f_A) de apariţie a unui eveniment A pe parcursul repetării de n ori a unui experiment este dată de relaţia:

$$f_A(n) = N_A(n)/n \qquad (4.1.1)$$

unde $N_A(n)$ este numărul de apariţii ale evenimentului A pe parcursul celor n repetări (**frecvenţa absolută** a evenimentului A).

Aplicaţia 4.1.2

Demonstraţi următoarele proprietăţi ale frecvenţei relative:

i) $0 \le f_A(n) \le 1$, $\forall$ eveniment A;

ii) $\sum_i f_{A_i}(n) = 1$, cu evenimente A_i disjuncte şi formând o partiţie a spaţiului S;

iii) $f_C(n) = f_A(n) + f_B(n)$, cu $C = A \cup B$ şi $A \cap B = \varnothing$.

* * *

În condiţii de regularitate statistică, **raportul pe termen lung** (pentru un număr foarte mare de repetări), exprimat de frecvenţa relativă, tinde către o constantă (notată p_A) numită **probabilitate** de apariţie a evenimentului A:

$$\lim_{n \to \infty} f_A(n) = p_A \qquad (4.1.2)$$

4.2 Axiomele probabilităţii

Utilizarea relaţiei (4.1.2) ca definiţie a probabilităţii generează dificultăţi în dezvoltarea unei teorii a probabilităţilor, deoarece: nu este clar în ce sens matematic există limita; un experiment nu poate fi repetat la infinit; prezintă inconsistenţă în cazul evenimentelor cu apariţie rară. De aceea, are sens dezvoltarea unei teorii abstracte a probabilităţilor, plecând de la premisele:

I. Un experiment aleator a fost definit prin procedura de desfăşurare şi prin măsurătorile şi observaţiile ce se efectuează, iar spaţiul de eşantionare S a fost identificat.

II. S-a specificat (stabilit) o familie A de evenimente, astfel încât:

$$\mathrm{S} \in \mathrm{A} \text{ şi } \varnothing \in \mathrm{A};$$

$$\forall A \in \mathrm{A} \text{ şi } A^C \in \mathrm{A}; \qquad (4.2.1)$$

$$A_{n \in \mathrm{N}} \in \mathrm{A} \rightarrow \bigcup_{n \in \mathrm{N}} A_n \in \mathrm{A}$$

(A este un *tribut* pe S, iar cuplul (S, A) se numeşte *spaţiu măsurabil*), [Ciuc83].

III. Fiecărui eveniment, A, îi este alocat (ataşat) un număr numit **probabilitate** şi notat $P[A]$, astfel încât următoarele axiome sunt satisfăcute:

Axioma 1: $0 \le P[A]$;

Axioma 2: $P[\mathrm{S}] = 1$;

Axioma 3: Fie A şi B două evenimente astfel încât $A \cap B = \varnothing$, atunci:

$$P[A \cup B] = P[A] + P[B] \qquad (4.2.2)$$

Axioma 3'. Fie $A_1, A_2, \cdots$ o secvenţă de evenimente, astfel încât:

$A_i \cap A_j = \varnothing$, pentru $\forall i \ne j$, atunci $P\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} P[A_k]$ (4.2.3)

Ca *măsură* a probabilităţii, $P[*]$ este o aplicaţie a lui A în □, iar tripletul $(\mathrm{S}, \mathrm{A}, P)$ se numeşte **spaţiu probabilizat** [Ciuc 83].

Un **model matematic** este **probabilistic** atunci când condiţiile de efectuare a experimentului vizat (ce prezintă regularitate statistică) determină exact *probabilităţile* de apariţie a realizărilor sale. Construirea unui model probabilistic constă în particularizarea premiselor **I**, **II** şi **III** la cazul studiat (situaţia reală). Domeniile de aplicabilitate sunt: sistemele de transmisiuni, comunicaţii prin medii zgomotoase, prelucrarea semnalelor, sisteme cu resurse partajate, sisteme de servire, fiabilitatea şi disponibilitatea sistemelor etc. Ca orice modelare, calculul probabilităţilor este o tehnică matematică de descriere a fenomenelor aleatorii. Validitatea sa într-o aplicaţie concretă nu poate fi înfăptuită decât cu ajutorul metodelor statistice de interpretare a observaţiilor.

Regula prin care unui eveniment A al unui experiment E i se alocă (ataşează) o probabilitate $P[A]$ se numeşte **lege de probabilitate.** În cazul unui spaţiu de eşantionare discret, legea poate fi specificată alocând probabilităţi evenimentelor elementare, iar în cazul continuu dând probabilităţi intervalelor.

Aplicaţia 4.2.1

Fie un experiment aleatoriu E, cu spaţiul de eşantionare discret $\mathrm{S}_E = \{s_1, s_2, s_3, \cdots, s_n\}$. Considerând evenimente de forma $A = \{s_{A,1}, s_{A,2}, \cdots, s_{A,j}\}$, să se arate că relaţia $P[A] = j/n$ reprezintă o lege de probabilitate, unde *j* este numărul de realizări ce îndeplinesc evenimentul A.

* * *

Aplicaţia 4.2.2

Durata de prelucrare a unui apel telefonic variază între 0 şi 1 secunde, [Nic95]. Considerând că toate realizările sunt echiprobabile, să se stabilească:

i) o lege de probabilitate potrivită aplicaţiei analizate;

ii) probabilitatea ca durata să se depărteze cu cel puţin 0,3 secunde de centrul intervalului.

Rezolvare: i) În acest caz, alocarea de probabilităţi unei realizări c, $(0 \le c \le 1)$ nu are sens, întrucât spaţiul realizărilor $\mathrm{S} = [0,1]$ este infinit nenumărabil şi deci:

$$P[\{c\}] = 0 .$$

Fie următoarea afirmaţie: "*probabilitatea ca o realizare să aparţină unui subinterval al mulţimii* S *este egală cu lungimea intervalului*", adică:

$$P[\,[a,b]\,] = b - a \text{ pentru } 0 \le a \le b \le 1$$

Aceasta este o lege de probabilitate, deoarece satisface axiomele (de verificat!).

ii) Evenimentul în discuţie are următoarea formă analitică:

$$A = [0;0,2] \cup [0,8;1] .$$

Intervalele fiind disjuncte putem aplica axioma 3, deci:

$$P[A] = P[0;0,2] + P[0,8;1] = 0,4$$

* * *

Aplicaţia 4.2.3

În urma măsurării timpilor de funcţionare ai unor componente electronice de acelaşi tip s-a ajuns la concluzia că: "*proporţia componentelor al căror timp de funcţionare depăşeşte durata* t *scade exponenţial cu rata* α". Să se determine legea de probabilitate care satisface acest rezultat.

Rezolvare: Spaţiul de eşantionare este: $\mathrm{S} = (0,\infty)$. Evenimentul conţinut în enunţ, $A = \{\text{timpul de funcţionare depăşeşte durata } t\}$ se exprimă analitic în forma $A = \{x : x > t\} = (t,\infty)$. Acestui eveniment i se alocă, conform afirmaţiei, următoarea valoare: $P[A] = P[(t,\infty)] = \exp(-\alpha t)$, pentru $t > 0$ şi $\alpha > 0$

Verificăm dacă această alocare corespunde unei legi probabilistice:

i) se verifică Axioma 1: $0 \le P[(t,\infty)] = \exp(-\alpha t)$ pentru $\forall t > 0$ şi $\alpha > 0$;

ii) se verifică Axioma 2: $P[\mathrm{S}] = P[(0,\infty)] = 1$;

iii) fie evenimentele: $B = (r,s]$ şi $C = (s,\infty)$.

Conform Axiomei 3: $P[(r,\infty)] = P[(r,s]] + P[(s,\infty)]$, deci:

$$P[(r,s]] = P[(r,\infty)] - P[(s,\infty)] = \mathrm{e}^{-\alpha \cdot r} - \mathrm{e}^{-\alpha \cdot s}$$

* * *

Deoarece orice alocare ce respectă axiomele este posibilă, rămâne în sarcina utilizatorului teoriei de a determina legea de probabilitate, precum şi modul de interpretare a probabilităţilor. Rezultatele teoretice corespunzătoare modelului adoptat, trebuie să poată fi interpretate, în cazul aplicaţiilor reale, ca rapoarte pe termen lung al unor mărimi de interes (de exemplu: frecvenţa relativă, media).

Legătura dintre această mărime abstractă, ***probabilitatea***, şi mărimea intuitivă ***frecvenţa relativă*** este oferită de două teoreme importante:

legea (slabă a) numerelor mari şi **teorema limită centrală** (tratate în detaliu în subcapitolele 3.3.5 - 3.3.7).

Legea numerelor mari arată că probabilitatea ca raportul $N_A(n)/n$ să se apropie cât dorim de mult de p_A, tinde către 1 când $n \to \infty$. Ne este permis, astfel, să interpretăm p_A ca o frecvenţă relativă.

Teorema limită centrală "măsoară" abaterea dintre modelul ales (în cazul de faţă, valoarea p_A) şi frecvenţa relativă $f_A(n)$ obţinută în urma unui număr finit de încercări.

În calculul probabilităţilor sunt utile următoarele **proprietăţi** (relaţii):

1. $P[A^C] = 1 - P[A]$, unde A^C este evenimentul complementar evenimentului *A*;
2. $P[A] \le 1$ - utilă în verificarea corectitudinii calculelor (vezi şi axioma 1);
3. $P[\varnothing] = 0$;
4. Dacă $A_1, A_2, \cdots, A_n$ sunt evenimente disjuncte atunci:
$$P\left[\bigcup_{k=1}^{n} A_k\right] = \sum_{k=1}^{n} P[A_k], \; n \ge 2$$
(utilă în calculul probabilităţilor unor evenimente complexe);
5. $P[A \cup B] = P[A] + P[B] - P[A \cap B]$ şi $P[A \cup B] < P[A] + P[B]$;
6. $P\left[\bigcup_{k=1}^{n} A_k\right] = \sum_{j=1}^{n} P[A_j] - \sum_{j \ne k} \sum_{j \ne k} P[A_j \cap A_k] + \ldots + (-1)^{n+1} P[A_1 \cap A_2 \cap \ldots \cap A_n]$;
7. $P\left[\bigcup_{k=1}^{n} A_k\right] \le \sum_{k=1}^{n} P[A_k]$ (utilă în obţinerea de margini superioare pentru probabilităţile de interes);
8. Dacă $A \subset B$ atunci: $P[A] \le P[B]$ (utilă în obţinerea de margini superioare pentru probabilităţile de interes).

4.3 Probabilităţi condiţionate

❖ **Probabilitatea condiţionată** ca evenimentul A să se îndeplinească dacă evenimentul B a avut loc este definită de relaţia:

$$P[A \mid B] = P[A \cap B] / P[B] \qquad (4.3.1)$$

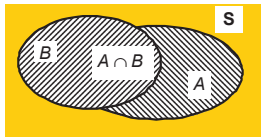


Figura 4.3.1: Diagrama Venn a intersecţiei mulţimilor *A* şi *B*

Cu alte cuvinte, conform figurii 4.3.1, dacă o realizare ζ a experimentului satisface evenimentul B, spaţiul realizărilor, în care se urmăreşte dacă a fost îndeplinit şi evenimentul A, se reduce la mulţimea B. În această situaţie, evenimentul A are loc dacă şi numai dacă $\zeta \in A \cap B$.

Aplicaţia 4.3.1

Pentru experimentul "*aruncarea cu zarul*", să se determine probabilitatea evenimentelor:

$$A=\{\textit{faţa } 2\},\ B=\{\textit{faţa } 1 \text{ sau } \textit{faţa } 2\},\ C=\{\textit{o faţă impară}\},$$

precum şi probabilităţile condiţionate: $P[A|C]$ şi $P[B \mid C]$.

* * *

Aplicaţia 4.3.2

La un nod de comutaţie, [Bor94], se recepţionează mesaje ce sunt distribuite echiprobabil unui număr de N terminale identificate prin adresele $0,1,2,\cdots,N$. Considerând următoarele evenimente:

$A=\{\text{mesajul recepţionat se transmite unui terminal de adresă pară}\}$

$B=\{\text{mesajul recepţionat se transmite unui terminal de adresă impară}\}$

$C=\{\text{mesajul recepţionat se transmite unui terminal de adresă} \le M, \text{ cu } M<N\}$

i) Să se precizeze semnificaţiile probabilităţilor: $P[B|A]$ şi $P[C|A]$;

ii) Să se determine expresiile lor. Caz particular: $N=25$ şi $M=16$.

Indicaţie: Fără a restrânge generalitatea, considerăm N şi M numere pare.

♦ *Varianta intuitivă:* Înaintea efectuării experimentului, spaţiul realizărilor este: $S=\{0,1,\cdots,N\}$. După efectuarea experimentului **se ştie** că evenimentul A s-a îndeplinit, deci **spaţiul redus** al realizărilor pe care se pot înfăptui evenimentele C sau B este: $S_r=A=\{0,2,4,\cdots,N-4,N-2\}$. Deoarece evenimentul *B* conţine doar realizări impare rezultă că $P[B \mid A]=0$.

În privinţa evenimentului C, spaţiul redus conţine realizările ce îl satisfac, şi anume: $\{0,2,4,\cdots,M\}$. Conform ipotezei, realizările sunt echiprobabile, deci se calculează probabilitatea $P[C|A]$ ca raport dintre:

$\{\text{numărul realizărilor din } S_r \text{ ce satisfac } C\}$, şi

$\{\text{numărul total al realizărilor din } S_r\}$, adică:

$$P[C \mid A]=(M/2+1)/(N/2)=(M+2)/N=18/25=72\%$$

♦ *Varianta analitică:* $P[B|A]=P[B\cap A]/P[A]=P[\varnothing]/P[A]=0$.

Deoarece:

- $P[C\cap A]$ este egală cu raportul dintre $\{\text{numărul realizărilor pare din } S_r \text{ mai mici sau egale cu } M\}$ şi $\{\text{numărul realizărilor din } S_r\}$, adică:

$$P[C\cap A]=(M/2+1)/N,$$

- $P[A]$ este egală cu raportul dintre $\{$numărul realizărilor pare din $S_r\}$ şi $\{$numărul realizărilor din $S_r\}$, adică: $P[A]=\dfrac{N/2}{N}$. Atunci rezultă că:

$$P[C \mid A]=\frac{P[C\cap A]}{P[A]}=\frac{M/2+1}{N}\cdot\frac{N}{N/2}=\frac{M+2}{N}.$$

* * *

❖ **Teorema probabilităţii totale**: Fie $B_1,B_2,\cdots,B_n$ o partiţie a spaţiului de eşantionare S şi A un eveniment oarecare. Probabilitatea de apariţie a evenimentului A este dată de relaţia:

$$P[A]=P[A|B_1]\cdot P[B_1]+P[A|B_2]\cdot P[B_2]+\ldots+P[A|B_n]\cdot P[B_n] \quad (4.3.2)$$

Această teoremă, împreună cu relaţia matematică ce îi corespunde, este utilă atunci când experimentul poate fi văzut ca secvenţă de mai multe subexperimente.

Aplicaţia 4.3.3

Abonaţii unui nod de comutaţie sunt împărţiţi **după durata** serviciilor cerute în două categorii. Pentru prima categorie, durata de servire urmează o lege exponenţială de parametru 1000μ, iar pentru a doua, aceeaşi lege de parametru μ. Să se determine probabilitatea ca un serviciu oarecare să se desfăşoare **încă** după *t* secunde de la iniţierea lui, ştiind că:

(a) cererea de servicii este aceeaşi pentru toţi abonaţii nodului de comutaţie;

(b) numărul abonaţilor din prima categorie este de patru ori mai mare decât al abonaţilor din a doua categorie.

Caz particular: $\mu=1$ serviciu/secundă şi $t=10$ secunde.

Indicaţie: Pentru rezolvarea aplicaţiei prezintă interes următoarele evenimente:

$A=\{\text{serviciul observat aparţine primei categorii}\}$;

$B=\{\text{serviciul observat aparţine categoriei a doua}\}$;

$C=\{\text{serviciul se desfăşoară încă după } t \text{ secunde de la iniţierea sa}\}$.

Evenimentele A şi B realizează o partiţie după categorii a spaţiului realizărilor, deci: $S=A\cup B$ unde $A\cap B=\varnothing$. Aplicând teorema probabilităţii totale obţinem: $P[C]=P[C|A]\cdot P[A]+P[C|B]\cdot P[B]$.

Pentru a determina $P[A]$ şi $P[B]$ ţinem cont de ipotezele (a) şi (b) din care deducem că cererile de servicii sunt echiprobabile.

Dacă n este numărul total al abonaţilor, iar n_1 şi n_2 sunt numerele de abonaţi din cele două categorii, atunci:

$P[A]=n_1/n=4/5$ şi $P[B]=n_2/n=1/5$.

Ştiind din datele aplicaţiei că: $P[C|A]=\exp(-1000\mu t)$ şi $P[C|B]=\exp(-\mu t)$, atunci rezultă: $P[C]=4/5\cdot\exp(-1000\mu t)+1/5\cdot\exp(-\mu t)=...$ etc.

* * *

❖ **Regula Bayes**: Fie $B_1,B_2,\cdots,B_n$ o partiţie lui S şi A un eveniment. Probabilitatea realizării evenimentului B_j când a avut loc evenimentul A este:

$$P[B_j|A]=\frac{P[A\cap B_j]}{P[A]}=\frac{P[A|B_j]\cdot P[B_j]}{\sum_{k=1}^{n}P[A|B_k]\cdot P[B_k]} \qquad (4.3.3)$$

Această regulă este utilă când evenimentele B_j sunt caracterizate prin **probabilităţi a priori**, $P[B_j]$, adică înainte ca experimentul să aibă loc, şi prin **probabilităţi a posteriori**, $P[B_j|A]$, când, în urma experimentului, s-a realizat evenimentul A.

Aplicaţia 4.3.4

O sursă binară generează simbolurile 1 şi 0 cu probabilitatea p, respectiv $1-p$. La recepţie, decizia este perturbată de zgomot, probabilitatea detectării incorecte a unui simbol fiind ε, [Spăt83].

Două evenimente diferite sunt avute în vedere, şi anume:

$A=\{\text{s-a generat simbolul } i,(i=0,1)\}$, $B=\{\text{s-a recepţionat simbolul } j,(j=0,1)\}$.

i) Să se determine probabilitatea $P[A_i\ \&\ B_j]$, pentru $i=0,1$ şi $j=0,1$.

Observaţie: Simbolul $\&$, echivalent cu $\cap$, este mai general, aplicându-se şi în cazul unor evenimente cu realizări din spaţii de eşantionare diferite.

ii) Considerând probabilitatea $p=0,5$ să se afle care simbol este mai probabil să fi fost transmis dacă la recepţie s-a luat decizia 4.

Rezolvare:

i) Diagrama arbore a experimentului este prezentată în figura 4.3.2 şi, parcurgând-o, rezultă că: $P[A_0\ \&\ B_0]=(1-p)\cdot(1-\varepsilon)$; $P[A_0\ \&\ B_1]=(1-p)\cdot\varepsilon$;

$P[A_1\ \&\ B_0]=p\cdot\varepsilon$ şi $P[A_1\ \&\ B_1]=p\cdot(1-\varepsilon)$.

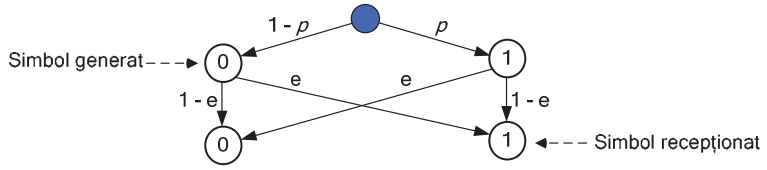


Figura 4.3.2: Diagrama arbore a realizărilor pereche în sistemul de transmisie binar

ii) Pentru a răspunde problemei trebuie să comparăm probabilităţile $P[A_0|B_1]$ şi $P[A_1|B_1]$. Regula Bayes oferă următoarele relaţii:

$$P[A_0|B_1]=\frac{P[A_0\ \&\ B_1]}{P[B_1]}=\frac{P[B_1|A_0]\cdot P[A_0]}{P[B_1]}$$

$$P[A_1|B_1]=\frac{P[A_1\ \&\ B_1]}{P[B_1]}=\frac{P[B_1|A_1]\cdot P[A_1]}{P[B_1]}$$

Considerând partiţia (A_0,A_1) a spaţiului de eşantionare, probabilitatea de la numitor se găseşte conform teoremei probabilităţii totale:

$$P[B_1]=P[B_1|A_0]\cdot P[A_0]+P[B_1|A_1]\cdot P[A_1]=\varepsilon\cdot(1-p)+(1-\varepsilon)\cdot p\big|_{p=0,5}=0,5$$

Deci: $P[A_0|B_1]=\frac{\varepsilon\cdot 0,5}{0,5}=\varepsilon$ şi $P[A_1|B_1]=\frac{(1-\varepsilon)\cdot 0,5}{0,5}=1-\varepsilon$.

Interpretarea rezultatului:

- pentru $\varepsilon>0,5$ este mai probabil să fi fost emis simbolul "0";
- pentru $\varepsilon<0,5$ este mai probabil să fi fost emis simbolul "1" (cazul real);
- pentru $\varepsilon=0,5$ simbolurile emise sunt "la fel de probabile".

4.4 Independenţa evenimentelor

Două evenimente *A* şi *B* sunt **independente** dacă:

$$P[A\cap B]=P[A]\cdot P[B] \qquad (4.4.1)$$

Rezultă deci că:

$$P[A|B]=P[A] \text{ şi } P[B|A]=P[B] \qquad (4.4.2)$$

Aplicaţia 4.4.1

Fie un concentrator cu o singură linie de ieşire, LI, la care sunt conectaţi 4 abonaţi identificaţi prin perechile $(1,1)$, $(2,1)$, $(3,2)$ şi $(4,2)$. Primul element din pereche reprezintă adresa, iar al doilea categoria abonatului. Verificaţi independenţa următoarelor evenimente luate două câte două:

$A=\{\text{un abonat din prima categorie are acces la LI}\}$;

$B=\{\text{un abonat de adresă pară are acces la LI}\}$;

$C=\{\text{un abonat din a doua categorie are acces la LI}\}$.

Se va considera că accesul la linia de ieşire LI este distribuit echiprobabil între abonaţi.

Rezolvare: Scrierea analitică a evenimentelor este următoarea:

$A=\{(1,1);(2,1)\}$, $B=\{(2,1);(4,2)\}$, $C=\{(3,2);(4,2)\}$

Datorită distribuţiei echiprobabile a accesului la linia de ieşire, înseamnă că:

$$P[A]=P[B]=P[C]=2/4=1/2$$

Pentru evenimentele: $A\cap B=\{(2,1)\}$, $A\cap C=\varnothing$ şi $B\cap C=\{(4,2)\}$ avem probabilităţile: $P[A\cap B]=1/4$, $P[A\cap C]=0$ şi $P[B\cap C]=1/4$. Deci:

- $\underbrace{P[A\cap B]}_{1/4}=\underbrace{P[A]}_{1/2}\cdot\underbrace{P[B]}_{1/2}$ adică evenimentele sunt independente. Cu alte cuvinte, realizarea evenimentului B nu modifică probabilitatea evenimentului $P[A|B]=P[A]$ şi viceversa.
- $\underbrace{P[A\cap C]}_{0}\neq\underbrace{P[A]}_{1/2}\cdot\underbrace{P[C]}_{1/2}$ adică evenimentele **nu** sunt independente; realizarea evenimentului C face **certă** nerealizarea evenimentului A ($P[A|C]=0$) şi viceversa.
- $\underbrace{P[B\cap C]}_{1/4}=\underbrace{P[B]}_{1/2}\cdot\underbrace{P[C]}_{1/2}$, adică evenimentele sunt independente: realizarea evenimentului C nu modifică probabilitatea evenimentului B ($P[B|C]=P[B]$) şi viceversa.

* * *

În general, evenimentele $A_1,A_2,\cdots,A_n$ sunt independente dacă pentru orice combinaţie de elemente şi oricâte dintre ele există relaţia:

$$P[A_i\ \&\cdots\&\ A_j\ \&\cdots\&\ A_k]=P[A_i]\cdot\ldots\cdot P[A_j]\cdot\ldots\cdot P[A_k] \quad (4.4.3)$$

Aplicaţia 4.4.2

Aflaţi numărul de relaţii ce definesc independenţa unui număr de n evenimente.

Răspuns: 2^n-n-1.

4.5 Experimente secvenţiale

O succesiune de experimente elementare $E_1,E_2,\cdots,E_n$, numite **subexperimente**, constituie un **experiment secvenţial**. Realizarea s a unui astfel de experiment este un vector ale cărui elemente s_k sunt realizările subexperimentelor implicate $(1,2,\cdots,k,\cdots,n)$: $s=(s_1,s_2,\cdots,s_k,\cdots,s_n)$. Spaţiul de eşantionare S este produsul cartezian al spaţiilor individuale: $S=S_1\times S_2\times\ldots\times S_n$.

Independenţa evenimentelor $A_1,A_2,\cdots,A_n$ ce apar într-un experiment secvenţial este indusă de independenţa subexperimentelor sale, dacă fiecare eveniment A_k depinde doar de realizările subexperimentului k. În acest caz:

$$P[A_1\ \&\ A_2\ \&\ \ldots\ \&\ A_n]=P[A_1]\cdot P[A_2]\cdot\ldots\cdot P[A_n].$$

Pe baza acestui raţionament, calculele laborioase, impuse de verificarea relaţiilor ce definesc independenţa evenimentelor, sunt evitate.

Secvenţele de subexperimente independente în care se urmăreşte numărul apariţiilor (*succeselor*) unui anumit eveniment (de exemplu: recepţia simbolului "1" dintr-o secvenţă binară) se numesc **încercări Bernoulli independente**. Ele sunt modelate cu ajutorul legii de probabilitate binomială, ce are următorul enunţ:

Fie k numărul de succese obţinute în n încercări Bernoulli independente. atunci probabilitatea lui k este dată de **legea de probabilitate binomială:**

$$p_n(k)=C_n^k\cdot p^k\cdot(1-p)^{n-k}\ \text{ pentru }\ k=\overline{0,n} \quad (4.5.1)$$

unde: $p_n(k)$ este probabilitatea apariţiei a k succese în n încercări,

$C_n^k=\dfrac{n!}{k!\cdot(n-k)!}=\dfrac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{1\cdot2\cdot\ldots\cdot k}$ este ***coeficientul binomial***,

p este probabilitatea apariţiei succesului într-o încercare.

În general, pentru evitarea dificultăţilor generate de termenii factoriali, în calculele computerizate se utilizează formula recursivă:

$$p_n(k+1)=\frac{(n-k)\,p}{(k+1)(1-p)}\cdot p_n(k) \quad (4.5.2)$$

Aplicaţia 4.5.1

Un nod de comutaţie deserveşte N terminale. Să se determine probabilitatea ca numărul de terminale active să fie mai mare ca M, dacă p este probabilitatea ca un terminal să fie activ.

Caz particular: $N=8$; $M=6$; $p=0,3$.

Rezolvare: Fie evenimentele: $A_k=\{k\ \text{terminale sunt active}\}$, $k=\overline{0,N}$. Atunci:

$$P[\{\text{număr de terminale active}>M\}]=\sum_{k=M+1}^{N}P[\{k\ \text{terminale sunt active}\}]$$

$$=\sum_{k=M+1}^{N}C_N^k\cdot p^k\cdot(1-p)^{N-k}=0,00259$$

* * *

Aplicaţia 4.5.2

Fie un sistem de telecomunicaţii care permite realizarea unui număr de N convorbiri simultane între două localităţi prin multiplexarea în timp pe un canal de comunicaţie. Semnalele vocale sunt eşantionate cu perioada T [μsec] şi cuantizate pe Q [biţi]. Secvenţele binare, obţinute sunt încapsulate în pachete de lungime constantă, ce corespunde la τ [msec] de segment de semnal audio.

Pentru o corectă redistribuire la recepţie, fiecărui pachet i se ataşează adresele sursei şi destinaţiei. Faptul că, în medie, $s\,\%$ din pachetele obţinute corespund perioadelor de tăcere din convorbiri, conduce la ideea că numărul maxim, M, de pachete transmise de sistem pe durata τ poate fi mai mic decât cel al convorbirilor simultane, $M < N$.

Considerând că clienţii acceptă pierderea pentru $q\,\%$ pachete active, să se construiască un model probabilistic pe baza căruia să se proiecteze un astfel de sistem.

Caz particular: $N = 48$, $q = 1$ şi $s = 66$.

Rezolvare: Specificarea experimentului: la fiecare τ msec se obţin N pachete active şi inactive. Putem considera realizare a experimentului evenimentul elementar A, ce reprezintă numărul pachetelor active, produse într-un interval. În acest caz, spaţiul realizărilor este: $S = \{0,1,..N\}$. Evenimente de interes sunt:

$*$ $\{A \le M\}$ - toate pachetele active sunt transmise;

$*$ $\{A > M\}$ - un număr de $A - M$ pachete active sunt pierdute.

Probabilităţile ataşate evenimentelor elementare sunt generate de legea de probabilitate binomială: $p_k = C_N^k\, p^k\, (1-p)^{N-k}$, unde $p = 1 - s/100$ este raport pe termen lung al pachetelor active din totalul de pachetelor generate de o sursă.

Mărimea p_k este valoarea către care converge frecvenţa relativă a evenimentului $\{A = k\}$ după o lungă secvenţă de repetări:

$$\lim_{n\to\infty} f_k(n) = \lim_{n\to\infty} N_k(n)/n = p_k \text{, pentru } 0 \le k \le N,$$

unde $N_k(n)$ este numărul de realizări ale evenimentului $\{A = k\}$ în cele n repetări. Procentul de pachete active pierdute este dat de raportul dintre media numărului de pachete pierdute şi media pachetelor generate:

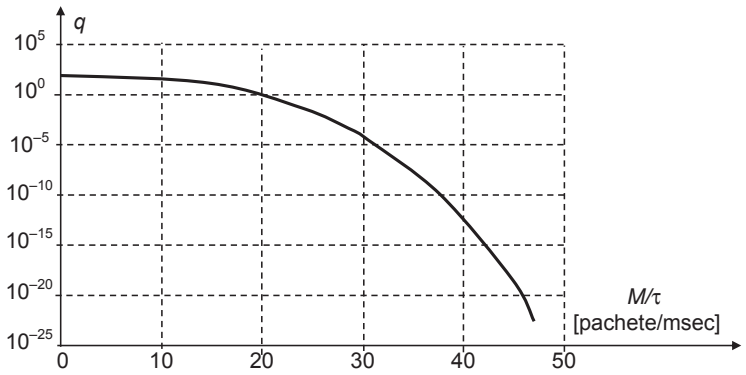


Figura 4.5.1: Variaţia cu rata de transmisie a procentului de pachete pierdute

$$\frac{q}{100} = \frac{\sum_{k=M+1}^{N}(k-M)\cdot N_k(n)}{\sum_{k=0}^{N} k\cdot N_k(n)} = \frac{\sum_{k=M+1}^{N}(k-M)\cdot p_k}{\sum_{k=0}^{N} k\cdot p_k}$$

Pentru cazul particular propus, se obţine curba din figura 4.5.1, din care rezultă că: $M = 20$ pachete.

* * *

Secvenţele de subexperimente independente în care se urmăreşte numărul succesului pentru mai multe evenimente reciproc exclusive sunt modelate cu ajutorul **legii de probabilitate multinomială**.

Fie $B_1, B_2, \cdots, B_M$ evenimente ce realizează o partiţie a spaţiului de eşantionare S. Probabilităţile de apariţie a lor în cadrul unui subexperiment, $P[B_j] = p_j$, satisfac relaţia: $p_1 + p_2 + \ldots + p_M = 1$.

Fie k_j numărul de succese ale evenimentului B_j, cu $j = \overline{1,M}$, obţinut în n repetiţii independente. Vectorul $(k_1, k_2, \cdots, k_M)$, ce specifică numărul succeselor pentru fiecare eveniment, satisface **legea de probabilitate multinomială:**

$$P[(k_1,k_2,...,k_M)] = \frac{n!}{k_1!\cdot k_2!\cdot \ldots \cdot k_M!}\prod_{j=1}^{M} p_j^{kj} \text{, cu } k_1 + k_2 + \ldots + k_M = n \quad (4.5.3.)$$

Aplicaţia 4.5.3

Un calculator multitasking, multiuser execută 3 categorii de programe caracterizate prin duratele: 10 msec, 100 msec şi 200 msec. Aceste programe sunt cerute de utilizatori cu probabilităţile: 0,2; 0,3 şi 0,5.

Să se determine probabilitatea execuţiei simultane a câte 3 programe din fiecare categorie.

Rezolvare: Mulţimea programelor de orice durată, S, se partiţionează în:

$B_1 = \{\text{mulţimea programelor de 10 msec}\}$, cu $P[B_1] = 0,2$;

$B_2 = \{\text{mulţimea programelor de 100 msec}\}$, cu $P[B_2] = 0,3$;

$B_3 = \{\text{mulţimea programelor de 200 msec}\}$, cu $P[B_3] = 0,5$;

Evenimentul $A = \{\text{se execută simultan câte 3 programe din fiecare categorie}\}$ are expresia analitică: $A = (3,3,3)$ şi aplicând legea de repartiţie multinomială obţinem: $P[(3,3,3)] = \frac{9!}{3!\cdot 3!\cdot 3!}\cdot (0,2)^3 \cdot (0,3)^3 \cdot (0,5)^3 = \ldots = 0,04536$

* * *

Secvenţele de subexperimente independente de tip Bernoulli în care se urmăreşte numărul de repetări până la apariţia primului succes sunt modelate cu ajutorul **legii geometrice de probabilitate**:

$$p(m) = P\left[A_1^C \,\&\, A_2^C \,\&\, \ldots \,\&\, A_{m-1}^C \,\&\, A_m \right] = p \cdot (1-p)^{m-1}, \text{ cu } m \geq 1 \quad (4.5.3)$$

unde: m este numărul de repetări până la apariţia succesului;

$A_1^C, A_2^C, \cdots, A_{m-1}^C$ este secvenţa de insuccese obţinute în primele m încercări;

A_m este succesul evenimentului în încercarea m ;

p este probabilitatea apariţiei succesului într-o încercare.

Probabilitatea ca numărul necesar de încercări, m, pentru obţinerea succesului să fie mai mare decât un număr fixat k este: $P[m > k] = (1-p)^k$.

Aplicaţia 4.5.4

Pentru realizarea unei comunicaţii printr-un mediu perturbator, mesajele sunt codate astfel încât la recepţie să se poată realiza procesul de detecţie al erorilor. În caz de eroare la recepţie, destinaţia cere sursei retransmisia mesajului în cauză.

Ştiind că mesajele se perturbă cu o probabilitate q, să se determine care este probabilitatea ca un mesaj să fie retransmis de cel puţin k ori. Caz particular: $q = 0{,}1$ şi $k = 2$.

Rezolvare: Considerând că *m* este numărul transmisiilor până când se obţine succesul (recepţia corectă a mesajului) avem:

$$P[m \geq k+1] = P[m > k] = q^k = 0{,}01$$

* * *

În cazul **lanţurilor** (*secvenţelor*) de subexperimente dependente, realizarea din cadrul unui subexperiment depinde de realizările anterioare. Astfel, o secvenţă particulară de realizări $s_0, s_1, \cdots, s_n$ are probabilitatea:

$$P[s_0 \,\&\, s_1 \,\&\, \ldots \,\&\, s_n] = \\ = P\left[s_n \middle| s_0 \,\&\, s_1 \ldots \,\&\, s_{n-1}\right] \cdot P\left[s_{n-1} \middle| s_0 \,\&\, \ldots \,\&\, s_{n-2}\right] \cdot \ldots \cdot P\left[s_1 \middle| s_0\right] \cdot P[s_0] \quad (4.5.4)$$

Lanţurile de subexperimente dependente la care realizarea curentă depinde numai de realizarea precedentă se numesc **lanţuri Markov** (de ordinul I). În acest caz, relaţia anterioară devine:

$$P[s_0 \,\&\, s_1 \,\&\, \ldots \,\&\, s_n] = P\left[s_n \middle| s_{n-1}\right] \cdot P\left[s_{n-1} \middle| s_{n-2}\right] \cdot \ldots \cdot P\left[s_1 \middle| s_0\right] \cdot P[s_0] \quad (4.5.5)$$

Aplicaţia 4.5.5

Semnalul codat numeric, obţinut de la o sursă este fragmentat în segmente vocale care sunt încapsulate în pachete de lungime constantă. Se consideră că pachetele sunt doar de două tipuri: 0, ce corespunzător perioadelor de tăcere din convorbire, şi 1, ce corespunzător activităţii sonore.

Pachetele tind să se succedă în salve de tip 0 cu probabilitatea p (după un simbol de tip 0 urmează tot un simbol de tip 0), sau de tip 1 cu probabilitatea

q (după un simbol de tip 1 urmează tot un simbol de tip 1). Se cere:

i) Imaginaţi un model capabil să descrie această situaţie.

ii) Determinaţi probabilitatea secvenţei iniţiale 001111110.

Caz particular: $p = 2/3$, $q = 1/4$.

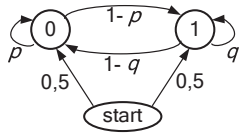


Figura 4.5.2: Diagrama unui lanţ Markov cu două realizări

Rezolvare:

i) Este vorba de un lanţ Markov, realizările unui subexperiment depinzând strict de realizările subexperimentului precedent. Modelul propus are diagrama din figura 4.5.2. Nodurile (stările) 0 şi 1 reprezintă o realizare posibilă a unui subexperiment, iar arcele conduc, cu probabilităţile menţionate, către o realizare posibilă din următorul subexperiment. "Start" reprezintă nodul de plecare al experimentul, iniţial sursa putându-se găsi, cu aceeaşi şansă, în starea de tăcere sau activitate sonoră.

ii) Probabilitatea cerută este dată de relaţia:

$$P[001111110] = P[0|1] \cdot \left\{P[1|1]\right\}^5 \cdot P[1|0] \cdot P[0|0] \cdot P[0] \\ = (1-q) \cdot q^5 \cdot (1-p) \cdot p \cdot 1/2$$

* * *

Aplicaţia 4.5.6

Reluaţi aplicaţia anterioară considerând că sursa este modelată prin intermediul unui lanţ Markov de ordinul II, adică în cazul în care realizarea curentă depinde numai de ultimele două realizări precedente.

Indicaţie: În această situaţie, stările diagramei vor preciza combinaţiile posibile ale ultimilor două realizări precedente.

* * *

Aplicaţia 4.5.7

Desenaţi diagrama corespunzătoare superpoziţiei a două surse independente de segmente vocale ce acţionează fiecare conform modelului din aplicaţia 4.5.4., precizând în dreptul fiecărui arc expresiile, funcţie de p şi q, ale probabilităţilor în cauză.

capitolul 5

Variabile aleatorii
- definiţii şi concepte de bază -

Variabila aleatorie X este o funcţie *fixă* şi *deterministă* care alocă (repartizează, distribuie) un număr real, $X(\zeta)$, fiecărei realizări ζ din spaţiul de eşantionare, S, al unui experiment aleator, E. Figura 5.1 prezintă un caz concret, cel al extragerii unei bile dintr-o urnă, pe lângă mărimile analitice precizate mai sus apărând şi **spaţiul de eşantionare** al variabilei, cu notaţia S_X.

Conform [Ciuc83], $X(\zeta)$ este o aplicaţie (măsurabilă) a spaţiului (S,A) pe mulţimea R înzestrată cu tributul lui Borel, B. Fie A un eveniment în spaţiul S şi B un eveniment în spaţiul S_X. Ele sunt **evenimente echivalente** dacă $A=\{\zeta : X(\zeta)\in B\}$. În acest caz:

$$P[B]=P[A]=P\left[\{\zeta : X(\zeta)\in B\}\right] \qquad (5.1)$$

Notaţia $P[*]$ din în relaţia anterioară nu reprezintă în toţi termenii acelaşi lucru, ci în ultimii doi este vorba de aplicaţia lui A pe R, iar în primul termen semnifică aplicaţia lui B pe R, ceea ce este de fapt o **lege de probabilitate**.

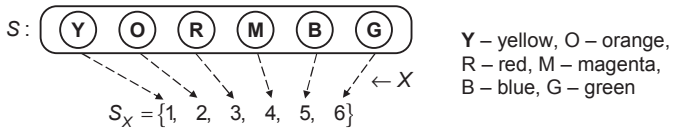


Figura 5.1: Experimentul de extragere a unei bile colorate şi mărimi analitice asociat

Comportamentul unei variabile aleatorii X este descris complet de funcţia de distribuţie, de funcţia densitate de probabilitate sau de probabilităţile elementare. Despre toate aceste concepte se fac precizările necesare înţelegerii lor, în cele ce urmează, însoţind rezultatele teoretice cu aplicaţii strâns legate de subiectul lucrării de faţă.

5.1. Funcţia de distribuţie

Funcţia de distribuţie, $F_X(x)$, a variabilei aleatorii X este probabilitatea apariţiei evenimentului $\{X\le x\}$:

$$F_X(x)=P\left[\{X\le x\}\right], \text{ pentru } -\infty<x<+\infty \qquad (5.1.1)$$

Proprietăţile funcţiei de distribuţie, ce rezultă de fapt din axiomele probabilităţii, sunt următoarele:

- $0\le F_X(x)\le 1$;
- $\lim_{x\to\infty} F_X(x)=1$;
- $\lim_{x\to-\infty} F_X(x)=0$;
- $F_X(a)<F_X(b)$ pentru orice $a<b$ (monoton crescătoare);
- $F_X(b)=\lim_{h\to 0 \text{ şi } h>0} F_X(b+h) \overset{\text{notat}}{=} F_X(b^*)$ (continuitatea la dreapta);
- $P[a<X\le b]=F_X(b)-F_X(a)$;
- $P[X=b]=F_X(b)-F_X(b^-)$ sau $P[X=b]=0$, dacă X este continuă în b;
- $P[a\le X\le b]=F_X(b)-F_X(a^-)$;
- $P[a<X<b]=F_X(b^-)-F_X(a)$.

O variabilă aleatoare X este **variabilă aleatorie continuă** dacă funcţia sa de distribuţie $F_X(x)$ este continuă peste tot şi poate fi definită ca o integrală dintr-o funcţie pozitivă $f(x)$:

$$F_X(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t \qquad (5.1.2)$$

O variabilă aleatorie X este **variabilă aleatorie discretă** dacă funcţia sa de distribuţie $F_X(x)$ este sub formă de scară, continuă la dreapta, având o mulţime măsurabilă (finită sau infinită) de puncte de salt $x_0,x_1,x_2,\cdots$ Aceste puncte reprezintă spaţiul de eşantionare S_X.

Funcţia de distribuţie a unei variabile aleatorii discrete se poate exprima ca o sumă ponderată de funcţii treaptă unitate:

$$F_X(x)=\sum_k p_X(x_k)\cdot u(x-x_k) \qquad (5.1.3)$$

unde: $p_X(x_k) = P[X = x_k]$ este **probabilitatea elementară** de apariţie a *evenimentului elementar* $\{X = x_k\}$.

O variabilă aleatorie X este **variabilă aleatorie mixtă** dacă funcţia sa de distribuţie $F_X(x)$ are, pe lângă o mulţime de salturi în punctele $x_0, x_1, x_2, \cdots$, cel puţin un interval pe care evoluează continuu.

Astfel de variabile aleatorii se pot genera în cadrul unor experimente secvenţiale dependente, în 2 trepte: o încercare Bernoulli este urmată de generarea variabilei aleatorii conform distribuţiei aleasă în subexperimentul precedent.

5.2 Funcţia de densitate de probabilitate

Funcţia de densitate de probabilitate, $f_X(x)$, a variabilei aleatorii X, dacă există, este derivata funcţiei de distribuţie $F_X(x)$:

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x} \qquad (5.2.1)$$

Funcţia $f_X(x)$ reprezintă o ***densitate de probabilitate*** întrucât probabilitatea ca variabila aleatorie X să apară în vecinătatea $(x, x+\mathrm{d}x)$ a unui punct x, este dată de relaţia (vezi figura 5.2.1): $P[x < X \le x + \mathrm{d}x] \approx f_X(x)\,\mathrm{d}x$

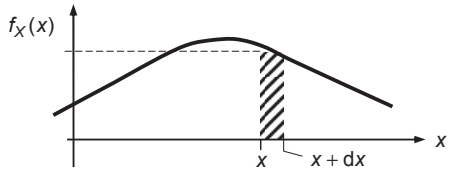


Figura 5.2.1: Specificarea probabilităţii unui interval infinitezimal prin intermediul funcţiei densitate de probabilitate

Orice funcţie $g(x)$, pozitivă, continuă pe subintervale, care respectă condiţia $\int_{-\infty}^{+\infty} g(x)\,\mathrm{d}x = C < \infty$, poate fi utilizată în construirea unei funcţii densitate de probabilitate, $f_X(x)$ pe baza relaţiei: $f_X(x) = g(x)/C$.

Deoarece $f_X(x)$ trebuie definită pe întreaga mulţime a numerelor reale, este necesar ca, pe eventualele intervale în care funcţia $g(x)$ nu ia valori, expresia densităţii de probabilitate să fie $f_X(x) = 0$.

În cazul variabilelor aleatorii discrete, cu funcţii de distribuţie discontinue în punctele $x_0, x_1, x_2, \cdots$, funcţia densitate de probabilitate va conţine funcţii delta

corespunzătoare punctelor de discontinuitate:

$$f_X(x) = \sum_k p_X(x_k) \cdot \delta(x - x_k) \qquad (5.2.2)$$

unde: $p_X(x_k) = P[X = x_k]$.

Funcţia de distribuţie corespunzătoare are expresia:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t = \sum_k p_X(x_k) \int_{-\infty}^{x} \delta(x - x_k)\,\mathrm{d}t = \sum_k p_X(x_k) \cdot u(x - x_k)$$

Pentru variabile aleatorii mixte, funcţia densitate de probabilitate va conţine atât funcţii delta, cât şi funcţii continue corespunzătoare.

Principalele proprietăţile ale funcţiei densitate de probabilitate sunt:

- $f_X(x) \ge 0$
- $P[a \le X \le b] = \int_a^b f_X(x)\,\mathrm{d}x$
- $F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$
- $\int_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x = 1$ (condiţia de normare).

Aplicaţia 5.2.1

Timpul de transmisie X a unui mesaj într-un sistem de comutaţie urmează o lege de probabilitate exponenţială, de parametru λ şi de forma $P[X > x] = \exp(-\lambda \cdot x)$, cu $x > 0$. Să se determine:

i) funcţia de distribuţie şi funcţia densitate de probabilitate;

ii) $P[T < X \le 2T]$, cu $T = 1/\lambda$.

Rezolvare:

i) $F_X(x) = P[X \le x] = 1 - P[X > x]$, deci:

$$F_X(x) = \begin{cases} 0 & \text{pentru } x < 0 \\ 1 - \exp(-\lambda \cdot x) & \text{pentru } x \ge 0 \end{cases}$$

$F_X(x)$ fiind o funcţie continuă, rezultă că:

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x} = \begin{cases} 0 & \text{pentru } x < 0 \\ \lambda \cdot \exp(-\lambda \cdot x) & \text{pentru } x \ge 0 \end{cases}$$

Graficele celor două funcţii sunt prezentate în figura 5.2.2.

ii) metoda 1: $P[T < X \le 2T] = P[T < X] - P[2T < X] = \mathrm{e}^{-1} - \mathrm{e}^{-2} \approx 0{,}233$;

metoda 2 : $P[T < X \le 2T] = F_X(2T) - F_X(T) = 1 - \mathrm{e}^{-2} - 1 + \mathrm{e}^{-1} \approx 0{,}233$;

metoda 3 : $P[T < X \le 2T] = \int_T^{2T} f_X(x)\,\mathrm{d}x = \mathrm{e}^{-\lambda \cdot x}\Big|_T^{2T} = \mathrm{e}^{-1} - \mathrm{e}^{-2} \approx 0{,}233$.

* * *

Aplicaţia 5.2.2

Timpul de aşteptare într-un sistem, W, este zero pentru un client dacă el

găseşte sistemul liber, iar în rest, adică dacă el găseşte sistemul ocupat, este distribuit exponenţial. Probabilitatea de a găsi sistemul liber este p. Găsiţi:

i) funcţia de distribuţie a variabilei aleatorii W;

ii) funcţia densitate de probabilitate.

Rezolvare: Dacă se utilizează teorema probabilităţii totale, se scrie că:

$$F_W(x) = P[W \le x] = P\left[W \le x \,\middle|\, \text{sistemul este liber}\right] \cdot p + P\left[W \le x \,\middle|\, \text{sistemul este ocupat}\right] \cdot (1-p)$$

Deoarece:

$$P\left[W \le x \,\middle|\, \text{sistem liber}\right] = \begin{cases} 0 & \text{pentru } x < 0 \\ & \text{(nu poate exista timp de aşteptare negativ)} \\ 1 & \text{pentru } x \ge 0 \ \ (W = 0 \le x) \end{cases}$$

$$P\left[W \le x \,\middle|\, \text{sistem ocupat}\right] = \begin{cases} 0 & \text{pentru } x < 0 \\ 1 - \exp(-\lambda \cdot x) & \text{pentru } x \ge 0 \end{cases},$$

rezultă deci că: $F_W(x) = \begin{cases} 0 & \text{pentru } x < 0 \\ p + (1-p)(1 - \mathrm{e}^{-\lambda x}) & \text{pentru } x \ge 0 \end{cases}$

Graficele celor două funcţii sunt prezentate în figura 5.2.3, în care săgeata reprezintă impulsul Dirac.

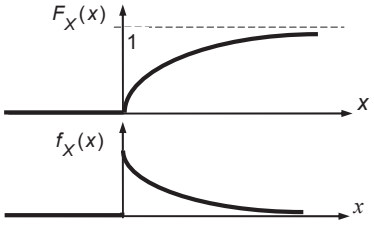


Figura 2.2.2: Funcţia de distribuţie şi funcţia densitate de probabilitate în cazul legii exponenţiale

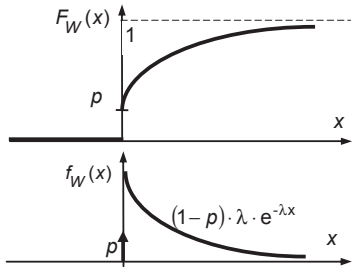


Figura 2.2.3: Funcţia de distribuţie şi funcţia densitate de probabilitate în cazul variabilei aleatorie mixte din aplicaţia 2.2.2

* * *

Funcţia de distribuţie condiţionată de un eveniment A, a variabilei aleatorii X este:

$$F_X(x \mid A) = \frac{P\left[\{X \le x\} \,\&\, A\right]}{P[A]}, \text{ dacă } P[A] > 0 \qquad (5.2.3)$$

Funcţia de densitate de probabilitate condiţionată de un eveniment A, a variabilei aleatorii X este (dacă derivata există!):

$$f_X(x \mid A) = \mathrm{d}F_X(x \mid A) / \mathrm{d}x \qquad (5.2.4)$$

Aplicaţia 5.2.3

Timpul de servire, X, al unui client urmează o funcţie de distribuţie continuă $F_X(x)$. Determinaţi funcţia de distribuţie condiţionată şi funcţia densitate de probabilitate condiţionată relativ la evenimentul:

$$A = \{\text{clientul este încă în servire la momentul } t\}.$$

Particularizaţi în cazul unei distribuţii exponenţiale pentru variabila X.

Rezolvare: Evenimentul $A = \{X > t\}$. Deci:

$$F_X(x \mid A) = P\left[X \le x \,\middle|\, X > t\right] = \frac{P\left[\{X \le x\} \,\&\, \{X > t\}\right]}{P\left[\{X > t\}\right]} = \begin{cases} 0 & \text{pentru } x < t \\ \dfrac{P[t < X \le x]}{1 - P[X \le t]} & \text{pentru } x \ge t \end{cases}$$

($F_X(x \mid A) = 0$ pentru $x < t$, deoarece este *imposibil* ca timpul de servire să fie sub x atunci când *se ştie* că el a depăşit valoarea t). Înlocuind probabilităţile cu expresii ce conţin funcţia de distribuţie, obţinem:

$$F_X(x \mid A) = \begin{cases} 0 & \text{pentru } x < t \\ \dfrac{F_X(x) - F_X(t)}{1 - F_X(t)} & \text{pentru } x \ge t \end{cases} \qquad (5.2.5)$$

$$f_X(x \mid A) = \frac{\mathrm{d}F_X(x \mid A)}{\mathrm{d}x} = \begin{cases} 0 & \text{pentru } x < t \\ \dfrac{f_X(x)}{1 - F_X(t)} & \text{pentru } x \ge t \end{cases} \qquad (5.2.6)$$

* * *

În procesul de modelare se poate recurge la o serie de distribuţii de variabile aleatorii. Descrierea completă, dar compactă, a celor mai uzuale distribuţii este prezentată în paragraful 5.6.

Aplicaţia 5.2.4

Să se demonstreze că, pentru n suficient de mari şi pentru p suficient de mici, distribuţia Poisson aproximează distribuţia binomială.

Demonstraţie: Facem notaţia $\alpha = n \cdot p$ (media numărului de succese). Probabilităţile generate de distribuţia binomială (vezi subcapitolul 1.5) au expresia: $p_k = \mathrm{C}_n^k \cdot p^k \cdot (1-p)^{n-k}$, pentru $k = \overline{1,n}$.

Relaţia de recurenţă care leagă aceste probabilităţi este:

$$\frac{p_{k+1}}{p_k} = \frac{\mathrm{C}_n^{k+1} p^{k+1} (1-p)^{n-k-1}}{\mathrm{C}_n^k p^k (1-p)^{n-k}} = \ldots = \frac{(1 - k/n) \cdot \alpha}{(k+1) \cdot (1 - \alpha/n)} \qquad (5.2.7)$$

Dar $p_{k+1}/p_k\big|_{n\to\infty} = \alpha/(k+1)$, deci $p_{k+1} = p_k \cdot \alpha/(k+1)$, pentru $k = \overline{1,n}$.

Pentru a putea folosi această relaţie, trebuie să determinăm una din probabilităţile p_k. În cazul de faţă: $p_0 = (1-p)^n = \left(1-\alpha/n\right)^n\big|_{n\to\infty} = \mathrm{e}^{-\alpha}$. În consecinţă, se obţine că: $p_1 = \mathrm{e}^{-\alpha}\,\alpha/1!$, $p_2 = \mathrm{e}^{-\alpha}\,\alpha^2/2!$, ..., $p_k = \mathrm{e}^{-\alpha}\,\alpha^k/k!$, ..., ceea ce corespunde unei distribuţii Poisson.

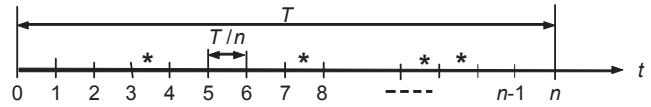


Figura 5.2.4: Desfăşurarea în timp a unei secvenţe de *n* încercări Bernoulli

Interpretare: O secvenţă de n încercări Bernoulli se poate desfăşura în spaţiu sau timp. În ultimul caz, ea poate fi reprezentată ca în figura 5.2.4, în care cele n încercări au loc într-un interval T, începând cu momentul 0, iar durata dintre două încercări succesive se presupune constantă şi egală cu T/n.

Simbolul * marchează în figura 5.2.4 apariţia succesului încercării din subintervalul corespunzător.

Distribuţia binomială ce descrie acest experiment oferă media numărului de succese, $n \cdot p$ în intervalul T. În acest caz, raportul $n \cdot p/T$ poate fi interpretat ca o rată medie de apariţie (sosire) a succeselor: $\lambda = n\,p/T = \alpha/T$.

Deci, variabila aleatorie N, ce reprezintă numărul de succese obţinute în n încercări Bernoulli desfăşurate în intervalul $[0,T]$ urmează o distribuţie binomială de parametri n şi p. Atunci când $n\to\infty$ şi $p\to 0$, N devine o variabilă aleatorie Poisson de parametru $\lambda = \alpha/T$.

* * *

Aplicaţia 5.2.5

Într-un canal de comunicaţii, probabilitatea ca un bit să fie eronat este de 10^{-3}. Determinaţi probabilitatea ca într-un bloc de 1000 de biţi să apară cel puţin 5 erori.

Rezolvare: notăm numărul biţilor eronaţi cu N şi probabilitatea de eroare cu p şi dorim să evaluăm: $P[N \ge 5] = 1 - P[N < 5]$.

- *Metoda* 1: aplicăm distribuţia binomială:

$$P[N \ge 5] = 1 - \sum_{k=0}^{4} \mathrm{C}_{1000}^{k}\, p^k (1-p)^{1000-k} = \cdots$$

- *Metoda* 2: aplicăm distribuţia Poisson, aceasta fiind permisă prin datele problemei:

$$P[N \ge 5] = 1 - \mathrm{e}^{-\alpha} \cdot \sum_{k=0}^{4} \alpha^k/k! = 1 - \mathrm{e}^{-1}\left(1 + 1/1! + 1/2! + 1/3! + 1/4!\right) = 0{,}00366$$

Aplicaţia 5.2.6

Să se demonstreze că distribuţia exponenţială este forma limită a distribuţiei geometrice.

Demonstraţie: Conform aplicaţiei 5.2.4, probabilitatea succesului unei încercări este dată de relaţia: $p = \alpha/n$.

Aşa cum este prezentat în figura 5.2.5, următoarele două evenimente:

- $\{N > k\}$
  $= \{\text{numărul } N \text{ de realizări până la primul succes să depăşească valoarea } k\}$
- $\{X > t\}$
  $= \{\text{durata } X \text{ până la obţinerea primului succes să depăşească valoarea } t\}$

sunt echivalente dacă se respectă relaţia: $k = \dfrac{t}{T/n}$.

Figura 5.2.5: Apariţia succesului după *t* unităţi de timp sau *k* repetări Bernoulli

În acest caz: $P\left[\{X > t\}\right] = P\left[\left\{N > \dfrac{t}{T/n}\right\}\right] = (1-p)^{n\cdot t/T} = \left[(1-\alpha/n)^n\right]^{t/T}$

La limită, când $n\to\infty$, rezultă că: $P[X > t] = \exp(-\alpha\, t/T)$, şi deci:

$$P[X \le t] = 1 - \exp(-\alpha \cdot t/T) = 1 - \exp(-\lambda \cdot t)$$

care este o distribuţie exponenţială de parametru λ (rata medie a apariţiilor).

* * *

Aplicaţia 5.2.7 **(lipsa memoriei variabilelor exponenţiale)**

Să se arate că pentru o variabilă exponenţială este adevărată relaţia $P[X > t+h \,|\, X > t] = P[X > h]$ şi să se interpreteze rezultatul.

Demonstraţie:

$$P[X > t+h \,|\, X > t] = \frac{P\left[\{X > t+h\} \,\&\, \{X > t\}\right]}{P\left[\{X > t\}\right]} = \frac{P[X > t+h]}{P[X > t]}$$

$$= \frac{\exp\left[-\lambda \cdot (t+h)\right]}{\exp(-\lambda \cdot t)} = \exp(-\lambda \cdot h) = P[X > h]$$

Interpretare: Probabilitatea ca realizarea unei variabile aleatorii exponenţiale să depăşească valoarea h este independentă de coordonata (timp, spaţiu etc.) de unde s-a început observarea (măsurarea) acesteia.

5.3 Funcţii de o variabilă aleatorie

Rezultatul generat de o funcţie deterministă şi reală, $g(x)$, care are ca argument valorile generate de o variabilă aleatorie, X, este tot o variabilă aleatorie, Y, numită **funcţie** (**transformare**) **de variabilă aleatorie**:

$$Y = g(X) \qquad (5.3.1)$$

Fie C un eveniment în spaţiul de eşantionare S_Y şi B evenimentul echivalent din S_X. Probabilităţile de apariţie a celor două evenimente sunt legate prin relaţia:

$$P[Y \in C] = P\left[g(X) \in C\right] = P[X \in B] \qquad (5.3.2)$$

Pentru a determina funcţia de distribuţie şi/sau funcţia de densitate spectrală a variabilei aleatorii Y se aplică relaţia precedentă, în care tipurile de evenimente din spaţiul S_X pot fi:

- $\{g(X) = y_k\}$, pentru a determina amplitudinea salturilor în acele puncte $(y_k, \cdots)$ atunci când este cunoscut că funcţia de distribuţie a lui Y are discontinuităţi;
- $\{g(X) \le y\}$, pentru a determina direct funcţia de distribuţie;
- $\{y < g(X) \le y + \mathrm{d}y\}$, pentru a determina funcţia densitate de probabilitate.

Aplicaţia 5.3.1

Determinaţi probabilităţile elementare ale variabilei aleatorii $Y = X^2$, dacă variabila aleatorie X:

i) are următoarele probabilităţi elementare:

$$p_X(-1) = 1/6;\ p_X(0) = 2/6;\ p_X(1) = 3/6.$$

ii) urmează o distribuţie Rayleigh.

Rezolvare:

ii) Ţinând cont că $\mathrm{S}_X = [0, \infty)$, înseamnă că transformării directă $y = x^2$, îi corespunde transformarea inversă $x = \sqrt{y}$. Prin urmare, evenimentului $(y, y + \mathrm{d}y)$, din spaţiul de eşantionare a variabilei Y, îi corespunde evenimentul $\left\{\sqrt{y}, \sqrt{y} + \left(2\sqrt{y}\right)^{-1}\mathrm{d}y\right\}$. Avem, deci, egalitatea: $f_Y(y)\mathrm{d}y = \dfrac{f_X(\sqrt{y})}{2\sqrt{y}} \cdot \mathrm{d}y$, de unde rezultă că: $f_Y(y) = \dfrac{1}{2\sigma^2}\exp\left(\dfrac{-y}{2\sigma^2}\right)$, adică Y este o variabilă exponenţială de parametru $1/(2\sigma^2)$.

5.4 Mărimi caracteristice ale unei variabile aleatorii

❖ **Valoarea aşteptată,** $\mathrm{E}[X]$**,** sau **media**, m, a unei variabile aleatorii X, notată şi prin simbolul $\overline{X}$, este dată de relaţiile**:**

$$\mathrm{E}[X] = \int_{-\infty}^{+\infty} t \cdot f_X(t)\mathrm{d}t, \text{ pentru cazul continuu} \qquad (5.4.1)$$

$$\mathrm{E}[X] = \sum_k x_k \cdot p_X(x_k), \text{ pentru cazul discret.} \qquad (5.4.2)$$

Pentru ca $\mathrm{E}[X]$ să existe, integrala (suma) trebuie să conveargă absolut:

$$\int_{-\infty}^{+\infty} |t| \cdot f_X(\mathrm{t})\mathrm{d}t < \infty \text{ sau } \sum_k |x_k| \cdot p_X(x_k) < \infty$$

În calcule, deseori sunt utile următoarele expresii ce sunt valabile pentru o variabilă aleatorie X pozitivă:

- continuă: $\mathrm{E}[X] = \int_0^{+\infty} \left(1 - F_X(t)\right)\mathrm{d}t$ (5.4.3)

- discretă: $\mathrm{E}[X] = \sum_{k=0}^{\infty} P[X > k]$ (5.4.4)

Valoarea aşteptată a unei variabile aleatorii $Y = g(X)$ se poate afla indirect, determinând funcţia densitate de probabilitate a lui Y sau direct cu ajutorul relaţiei:

$$\mathrm{E}[Y] = \mathrm{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) \cdot f_X(x)\mathrm{d}x \qquad (5.4.5)$$

Cele mai utile proprietăţi ale mediei sunt următoarele:

- $\mathrm{E}[C] = C$, unde C este o constantă oarecare,
- $\mathrm{E}[C \cdot X] = C \cdot \mathrm{E}[X]$,
- $\mathrm{E}[X + C] = \mathrm{E}[X] + C$
- $\mathrm{E}[Y] = \mathrm{E}\left[\sum_{k=1}^{n} g_k(X)\right] = \sum_{k=1}^{n} \mathrm{E}[g_k(X)]$

❖ **Deviaţia**, D, a variabilei aleatorii X faţă de medie este:

$$D = X - \mathrm{E}[X] \qquad (5.4.6)$$

❖ **Varianţa** (***dispersia***), σ^2, a variabilei aleatorii X, notată şi $\mathrm{VAR}[X]$, este:

$$\sigma^2 = \mathrm{VAR}[X] = \mathrm{E}\left[D^2\right] = \mathrm{E}\left[\left(X - \mathrm{E}[X]\right)^2\right] \qquad (5.4.7)$$

Cele mai utile proprietăţi ale varianţei sunt următoarele:

- $\mathrm{VAR}[C] = 0$, unde C este o constantă oarecare;
- $\mathrm{VAR}[X + C] = \mathrm{VAR}[X]$;

- $\mathrm{VAR}[C \cdot X] = C^2 \cdot \mathrm{VAR}[X]$;
- $\mathrm{VAR}[X] = \mathrm{E}[X^2] - \mathrm{E}^2[X]$.

❖ **Deviaţia standard** (***abaterea medie pătratică***) a variabilei aleatorii X este:

$$\sigma \overset{\text{notat}}{=} \mathrm{STD}[X] = \sqrt{\mathrm{VAR}[X]} \qquad (5.4.8)$$

❖ **Coeficientul de variaţie** este deviaţia standard raportată la medie, adică:

$$c_X = \frac{\mathrm{STD}[X]}{\mathrm{E}[X]} \qquad (5.4.9)$$

În rezolvarea unor probleme legate de ingineria traficului de telecomunicaţii se foloseşte **factorul de iregularitate**, definit ca raportul dintre varianţa şi media unei varibile, adică:

$$z = \frac{\mathrm{VAR}[X]}{\mathrm{E}[X]} \qquad (5.4.10)$$

❖ **Momentul de ordinul *n*** al variabilei aleatorii X este:

- pentru un spaţiu S_X continuu:

$$\mathrm{E}[X^n] = \int_{-\infty}^{+\infty} x^n f_X(x)\,\mathrm{d}x \qquad (5.4.11)$$

- pentru un spaţiu S_X discret:

$$\mathrm{E}[X^n] = \sum_k (x_k)^n \cdot p_X(x_k) \qquad (5.4.12)$$

❖ **Factorul de formă** al variabilei aleatorii X este:

$$\varepsilon_X = \mathrm{E}[X^2] \big/ \mathrm{E}^2[X] \qquad (5.4.13)$$

fiind legat de coeficientul de variaţie prin relaţia:

$$\varepsilon_X = 1 + c_X{}^2 \qquad (5.4.14)$$

Aceşti parametrii (media, deviaţia, varianţa, deviaţia standard, coeficientul de variaţie etc.) specifică **parţial** comportarea variabilei aleatorii căreia îi sunt asociaţi.

Aplicaţia 5.4.1

Determinaţi mărimile caracteristice ale variabilei Y, ce a fost considerată în aplicaţia 5.3.1.

* * *

Aplicaţia 5.4.2

Fie N numărul de cereri de mesaj generate de o unitate centrală şi transmise către un terminal până când acesta are un mesaj gata de transmis. Dacă presupunem că terminalul produce mesaje conform unei secvenţe de încercări Bernoulli independente, să se determine media variabilei aleatorii N.

Rezolvare: Conform ipotezei, variabila aleatorie N urmează o distribuţie geometrică de lege: $P[N=k] = p \cdot q^{k-1}$, unde $q = 1-p$

- varianta 1: $\mathrm{E}[N] = \sum_{k=1}^{+\infty} k \cdot p \cdot q^{k-1}$. Cum $\sum_{k=0}^{\infty} x^k = 1/(1-x)$ şi $\sum_{k=0}^{\infty} k \cdot x^{k-1} = 1/(1-x)^2$, rezultă că: $\mathrm{E}[N] = p/(1-q)^2 = 1/p$

- varianta 2: $\mathrm{E}[N] = \sum_{k=0}^{\infty} P[N > k]$

Cum $P[N>k] = 1 - P[N \le k] = 1 - (1-q^k) = q^k$, rezultă că:

$$\mathrm{E}[N] = \sum_{k=0}^{\infty} q^k = 1/p\,.$$

5.5 Transformatele unei variabile aleatorii

Cele trei metode de transformare, care urmează a fi prezentate în continuare, constituie un mijloc eficient de calcul al soluţiilor ecuaţiilor de tip integro-diferenţial, iar operaţia de înmulţire este mult mai rapidă (chiar şi pentru calculatoare) şi mai puţin predispusă erorilor decât operaţia de convoluţie.

➢ **Funcţia caracteristică** a unei variabile aleatorii X este:

- când spaţiul de eşantionare este continuu:

$$\Phi_X(\omega) = \mathrm{E}[\exp(j\omega X)] = \int_{-\infty}^{+\infty} f_X(x) \cdot \exp(j\omega x)\,\mathrm{d}x \qquad (5.5.1)$$

- când spaţiul de eşantionare este discret

$$\Phi_X(\omega) = \mathrm{E}[\exp(j\omega X)] = \sum_k p_X(x_k) \cdot \exp(j\omega x_k) \qquad (5.5.2)$$

Dacă spaţiul de eşantionare discret este mulţimea numerelor întregi, $\mathbb{Z}$ (cazul cel mai des întâlnit în practică), atunci:

$$\Phi_X(\omega) = \sum_k p_X(k) \cdot \exp(j\omega k) \qquad (5.5.3)$$

funcţia fiind periodică, de perioadă 2π.

Funcţia caracteristică, ca medie a funcţiei de o variabilă aleatorie $\exp(j\omega X)$, este practic transformata (seria) Fourier (excepţie făcând semnul exponenţialei) a funcţiei de densitate spectrală (a probabilităţilor evenimentelor elementare). Cunoscând expresia ei, putem aplica transformata Fourier inversă pentru a determina:

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi_X(\omega) \cdot \exp(j\omega X)\,\mathrm{d}\omega \qquad (5.5.4)$$

respectiv, când spaţiul de eşantionare este mulţimea $\mathbb{Z}$:

$$p_X(k) = \frac{1}{2\pi} \int_0^{2\pi} \Phi_X(\omega) \cdot \exp(-j\omega k)\,\mathrm{d}\omega \qquad (5.5.5)$$

Teorema momentelor:

$$\mathrm{E}\left[X^n\right] = \frac{1}{j^n} \cdot \left. \frac{\mathrm{d}^n \Phi_X(\omega)}{\mathrm{d}\omega^n} \right|_{\omega=0} \qquad (5.5.6)$$

Demonstraţie:

$$\Phi_X(\omega) = \mathrm{E}\left[\mathrm{e}^{j\omega X}\right] = \int_{-\infty}^{+\infty} \left\{1 + \frac{j\omega X}{1!} + \frac{(j\omega X)^2}{2!} + \ldots\right\} f_X(x)\mathrm{d}x$$

şi, considerând că toate momentele lui X sunt finite, seria poate fi integrată termen cu termen. Se obţine relaţia:

$$\Phi_X(\omega) = 1 + \mathrm{E}[X] \cdot \frac{j\omega}{1!} + \mathrm{E}\left[X^2\right] \cdot \frac{(j\omega)^2}{2!} + \ldots + \mathrm{E}\left[X^n\right] \cdot \frac{(j\omega)^n}{n!} + \ldots \qquad (5.5.7)$$

care, prin diferenţiere succesivă şi evaluare în $\omega = 0$, conduce la următoarele rezultate: $\mathrm{d}\Phi_X(\omega)/\mathrm{d}\omega\big|_{\omega=0} = \mathrm{E}[X]$,

... $\left. \mathrm{d}^n\Phi_X(\omega)/\mathrm{d}\omega^n \right|_{\omega=0} = j^n \cdot \mathrm{E}\left[X^n\right]$, q.e.d.

Aplicaţia 5.5.1

Determinaţi media şi varianţa unei variabile aleatorii ce urmează o distribuţie Gauss de parametrii μ şi σ utilizând funcţia caracteristică şi teorema momentelor.

Rezolvare:

$$\mathrm{E}[X] = \frac{1}{j} \cdot \left. \frac{\mathrm{d}\Phi_X(\omega)}{\mathrm{d}\omega} \right|_{\omega=0} = \frac{1}{j} \cdot (j\mu - \sigma^2\omega) \cdot \exp\left(j\mu\omega - \sigma^2\frac{\omega^2}{2}\right)\Big|_{\omega=0} = \mu$$

$$\mathrm{E}\left[X^2\right] = \frac{1}{j^2} \cdot \left. \frac{\mathrm{d}^2}{\mathrm{d}\omega^2}\Phi_X(\omega) \right|_{\omega=0} = -(j\mu - \sigma^2\omega - \sigma^2) \cdot \mathrm{e}^{j\mu\omega - \sigma^2\omega^2/2}\Big|_{\omega=0} = \sigma^2 + \mu^2$$

şi deci: $\mathrm{VAR}[X] = \mathrm{E}\left[X^2\right] - \mathrm{E}^2[X] = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$

* * *

De reţinut că: *funcţia densitate de probabilitate este complet specificată dacă toate momentele variabilei aleatorii sunt finite şi cunoscute*.

Aplicaţia 5.5.2

Să se determine funcţia densitate de probabilitate a variabilei aleatorii ale cărei momente respectă următoarea relaţie: $\mathrm{E}\left[X^n\right] = p$.

Rezolvare: Conform relaţiei (5.5.7), funcţia caracteristică a acestei variabile aleatorii este dată de expresia:

$$\Phi_X(\omega) = 1 + p \cdot j\omega + p \cdot (j\omega)^2/2! + \ldots = 1 - p + p \cdot \mathrm{e}^{j\omega}$$

Aplicând relaţia (5.5.4), obţinem: $f_X(x) = (1-p) \cdot \delta(x) + p \cdot \delta(x-1)$, adică, funcţia densitate de probabilitate a unei variabile aleatorii de tip Bernoulli.

Observaţie: expresia (5.5.7) se poate folosi şi atunci când se cunoaşte un număr finit de momente, caz în care relaţia (5.5.4) oferă o aproximare a funcţiei densitate de probabilitate.

* * *

- **Funcţia generatoare de probabilităţi**, $G_N(z)$, a unei variabile aleatorii N cu spaţiul de eşantionare în mulţimea numerelor naturale este:

$$G_N(z) = \mathrm{E}\left[z^N\right] = \sum_{k=0}^{\infty} p_N(k) \cdot z^k \qquad (5.5.8)$$

Al doilea termen reprezintă media unei funcţii de o variabilă aleatorie N, iar ultimul termen corespunde (cu excepţia semnului) unei transformate Z.

Numele acestei funcţii provine din faptul că, derivând-o succesiv, se obţin probabilităţile tuturor evenimentelor elementare:

$$p_N(k) = \frac{1}{k!} \cdot \left. \frac{\mathrm{d}^k G_N(z)}{\mathrm{d}z^k} \right|_{z=0} \qquad (5.5.9)$$

În mod asemănător, rezultă:

$$\mathrm{E}[N] = G_N'(1) \qquad (5.5.10)$$

$$\mathrm{VAR}[N] = G_N''(1) + G_N'(1) - \left(G_N'(1)\right)^2 \qquad (5.5.11)$$

La aceste formule se adăugă **relaţia de normare**:

$$G_N(1) = 1 \qquad (5.5.12)$$

Funcţia caracteristică a variabilei aleatorii N se obţine înlocuind variabila z cu $\exp(j\omega)$:

$$\Phi_N(\omega) = G_N\left(\exp(j\omega)\right) \qquad (5.5.13)$$

Aplicaţia 5.5.3

Să se demonstreze relaţiile (5.5.10) şi (5.5.11) şi să se aplice în cazul distribuţiei Poisson de parametru α.

Rezolvare:

- relaţia (5.5.10): $\mathrm{E}[N] = \sum_{k=0}^{\infty} k \cdot p_N(k) = \sum_{k=1}^{\infty} k \cdot p_N(k)$, dar:

$$\left. \frac{\mathrm{d}}{\mathrm{d}z} G_N(z) \right|_{z=1} = \sum_{k=1}^{\infty} k \cdot p_N(k) \cdot z^{k-1}\Big|_{z=1} = \sum_{k=1}^{\infty} k \cdot p_N(k) = \mathrm{E}[N]$$

- relaţia (5.5.11) conduce la:

$$\mathrm{VAR}[N] = \mathrm{E}\left[N^2\right] - \mathrm{E}^2[N], \text{ cu: } \mathrm{E}\left[N^2\right] = \sum_{k=1}^{\infty} k^2 \cdot p_N(k)$$

$$\left. \frac{\mathrm{d}^2}{\mathrm{d}z^2} G_N(z) \right|_{z=1} = \sum_{k=2}^{\infty} k \cdot (k-1) \cdot p_N(k) \cdot z^{k-2}\Big|_{z=1} = \sum_{k=2}^{\infty} k^2 \cdot p_N(k) - \sum_{k=2}^{\infty} k \cdot p_N(k)$$

$$= \sum_{k=1}^{\infty} k^2 \cdot p_N(k) - \sum_{k=1}^{\infty} k \cdot p_N(k)$$

Deci: $\mathrm{VAR}[N]=\left.\frac{\mathrm{d}^2G_N(z)}{\mathrm{d}z^2}\right|_{z=1}+\left.\frac{\mathrm{d}G_N(z)}{\mathrm{d}z}\right|_{z=1}-\left(\left.\frac{\mathrm{d}G_N(z)}{\mathrm{d}z}\right|_{z=1}\right)^2$

În cazul distribuţiei Poisson:

$$G_N(z)=\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}\cdot\frac{z^k}{\exp(\alpha)}=\frac{\mathrm{e}^{\alpha z}}{\mathrm{e}^{\alpha}}=\exp[\alpha\cdot(z-1)]$$

cu: $\left.\frac{\mathrm{d}G_N(z)}{\mathrm{d}z}\right|_{z=1}=\left.\alpha\cdot\mathrm{e}^{\alpha(z-1)}\right|_{z=1}=\alpha$ şi $\left.\frac{\mathrm{d}^2G_N(z)}{\mathrm{d}z^2}\right|_{z=1}=\left.\alpha^2\cdot\mathrm{e}^{\alpha(z-1)}\right|_{z=1}=\alpha^2$

Deci: $\mathrm{E}[N]=\alpha$ şi $\mathrm{VAR}[N]=\alpha^2+\alpha-\alpha^2=\alpha$

➢ **Transformata Laplace-Stieltjes LST** (***L**aplace-**S**tieltjes **T**ransform*), $\hat{X}(s)$, a funcţiei densitate de probabilitate, numită şi *funcţia generatoare de momente*, ce caracterizează o variabilă aleatorie X *pozitivă* şi *continuă*, este:

$$\hat{X}(s)=\mathrm{E}\left[\mathrm{e}^{-sX}\right]=\int_0^{\infty}f_X(x)\cdot\mathrm{e}^{-sx}\,\mathrm{d}x \qquad (5.5.14)$$

În cazul acestei transformate, **teorema momentelor** devine:

$$\mathrm{E}\left[X^n\right]=(-1)^n\left.\frac{\mathrm{d}^n\hat{X}(s)}{\mathrm{d}s^n}\right|_{s=0} \qquad (5.5.15)$$

Aplicaţia 5.5.4

Să se determine media şi varianţa unei variabile aleatorii X ce urmează o lege de distribuţie exponenţială de parametru λ utilizând transformata Laplace şi teorema momentelor.

Rezolvare:

$$\hat{X}(s)=\int_0^{\infty}f_X(x)\cdot\mathrm{e}^{-sx}\mathrm{d}x=\int_0^{\infty}\lambda\cdot\mathrm{e}^{-\lambda x}\cdot\mathrm{e}^{-sx}\mathrm{d}x=\lambda\cdot\int_0^{\infty}\mathrm{e}^{-(\lambda+s)x}\mathrm{d}x=\frac{\lambda}{\lambda+s};$$

$$\mathrm{E}[X]=(-1)\left.\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\lambda}{\lambda+s}\right)\right|_{s=0}=\frac{1}{\lambda}\ \text{şi}\ \mathrm{E}\left[X^2\right]=(-1)^2\left.\frac{\mathrm{d}^2}{\mathrm{d}s^2}\left(\frac{\lambda}{\lambda+s}\right)\right|_{s=0}=\frac{2}{\lambda^2}.$$

Deci: $\mathrm{VAR}[X]=\mathrm{E}\left[X^2\right]-\mathrm{E}^2[X]=2/\lambda^2-(1/\lambda)^2=1/\lambda^2$.

5.6 Distribuţii uzuale de variabile aleatorii

Căutarea unor modele probabilistice care să aproximeze satisfăcător diverse fenomene cu caracter aleatoriu a condus la identificarea a diverse distribuţii. Cele mai uzuale distribuţii folosite în domeniul telecomunicaţiilor sunt prezentate în continuare. Descrierea lor se referă la:

- condiţiile de valabilitate,
- spaţiul de eşantionare,
- funcţia densitate de probabilitate / funcţia de distribuţie,

- curbe reprezentative,
- valori aşteptate (media, varianţa),
- funcţia caracteristică,
- exemple de aplicabilitate.

5.6.1. Distribuţia Bernoulli

Se presupune că: variabila aleatorie X este **funcţie indicator** (I_A) pentru un eveniment A: $X=\begin{cases}1, \text{ dacă apare evenimentul } A\\ 0, \text{ în caz contrar}\end{cases}$

- Parametrii: $p=$ probabilitatea de apariţie a evenimentului A, $(0\le p\le 1)$.
- Probabilităţi elementare: $p_1=P[X=1]=p$ şi $p_0=P[X=0]=1-p$.
- Media: $\mathrm{E}[X]=p$;
- Varianţa: $\mathrm{VAR}[X]=p\cdot(1-p)$.
- Funcţia generatoare: $G_X(z)=1-p+p\cdot z$.
- Aplicaţie: modelarea mecanismului fundamental de generare a aleatoriului.

5.6.2. Distribuţia binomială

Se presupune că: variabila aleatorie X reprezintă numărul succeselor unui eveniment, înregistrate în n subexperimente (repetări) independente (este o sumă de n variabile aleatorii independente, identic distribuite după legea Bernoulli). Probabilitatea succesului în cadrul unei repetări este p.

- Spaţiul realizărilor: $S_X=\{0,1,2,\ldots,n\}$.
- Parametrii: $n\ (n>0)$ şi $p\ (0\le p\le 1)$
- Probabilităţi elementare:
  $p_X(k)=P[X=k]=\mathrm{C}_n^k\cdot p^k\cdot(1-p)^{n-k}$, pentru $k=\overline{0,n}$.
- Media: $\mathrm{E}[X]=n\,p$.
- Varianţa: $\mathrm{VAR}[X]=n\,p(1-p)$.
- Funcţia generatoare: $G_X(z)=(1-p+p\cdot z)^n$.
- Aplicaţii: numărul de biţi eronaţi într-un bloc de n biţi transmişi printr-un canal cu zgomot, numărul de surse active dintr-un număr n de surse etc.

5.6.3. Distribuţia geometrică

Versiunea I:

Se presupune că: variabila aleatorie X reprezintă numărul eşecurilor unui eveniment înregistrat pe parcursul unei secvenţe de subexperimente independente, tip Bernoulli, până la apariţia primului succes.

- Probabilitatea succesului în cadrul unei repetări este p.

- Spaţiul realizărilor : $S_X = \{0,1,2,\cdots\}$.
- Parametrii: p, cu $0 \le p \le 1$.
- Probabilităţi elementare: $p_X(k) = P[X = k] = p \cdot (1-p)^k$, cu $k \ge 0$
- Media: $E[X] = \frac{1-p}{p}$; Varianţa: $VAR[X] = \frac{1-p}{p^2}$.
- Funcţia generatoare: $G_X(z) = \frac{p}{1-(1-p)\cdot z}$.

***Versiunea II**:*

Se presupune că: variabila aleatorie X' este numărul de repetări până ce apare *primul succes* într-o secvenţă de subexperimente independente de tip Bernoulli.

- Spaţiul realizărilor: $S_{X'} = \{1,2,...\}$.
- Probabilităţi elementare: $p_{X'}(k) = p(1-p)^{k-1}$, $k = 1,2,\ldots$
- Media: $E[X'] = 1/p$; Varianţa: $VAR[X'] = (1-p)/p^2$.
- Funcţia generatoare: $G_{X'}(z) = \frac{p \cdot z}{1-(1-p)\cdot z}$.

Observaţii: i) $X' = X + 1$;

ii) **este singura variabilă aleatorie discretă ce deţine proprietate de lipsă a memoriei**.

- Aplicaţii:numărul de apeluri repetate până la ocuparea unei resurse; numărul de retransmisii într-un canal cu zgomot până la recepţionarea corectă; etc.

5.6.4. Distribuţia binomială negativă

Se presupune că: variabila aleatorie X este numărul de repetări până ce apar r succese într-o secvenţă de subexperimente independente de tip Bernoulli.

- Spaţiul realizărilor: $S_X = \{r, r+1, \cdots\}$.
- Probabilităţi elementare: $p_X(k) = C_{k-1}^{r-1} \cdot p^r \cdot (1-p)^{k-r}$, cu $k \ge r$
- Media: $E[X] = \frac{r}{p}$.
- Varianţa: $VAR[X] = \frac{r \cdot (1-p)}{p^2}$.
- Funcţia generatoare: $G_X(z) = \left\{ \frac{p \cdot z}{1-(1-p)\cdot z} \right\}^r$.

Observaţie: pentru $r = 1$ se obţine distribuţia geometrică *versiunea II*.

- Aplicaţii: asemănătoare cu cele de la distribuţia geometrică.

5.6.5. Distribuţia Poisson

Se presupune că:

i) variabila aleatorie X este numărul de apariţii ale unei anumite realizări pe parcursul unui interval de timp (ΔT);
ii) momentele apariţiilor sunt distincte;
iii) oricare interval finit conţine un număr finit de apariţii;
iv) oricare interval infinit conţine un număr infinit de apariţii;
v) evenimentele nu apar la momente predeterminate;
vi) numărul apariţiilor dintr-un interval este independent de numărul apariţiilor dintr-un alt interval disjunct .

- Spaţiul realizărilor: $S_X = \{0,1,2,\cdots\}$.
- Parametrii: $\alpha > 0$ (numărul mediu de evenimente ce apar în intervalul dat).
- Probabilităţi elementare: $p_X(k) = \alpha^k \cdot \exp(-\alpha)/k!$.
- Media: $E[X] = \alpha$; Varianţa: $VAR[X] = \alpha$.
- Funcţia generatoare: $G_X(z) = \exp[\alpha \cdot (z-1)]$.

Observaţii:

i) aproximează distribuţia binomială pentru p "mic" şi n "mare", considerând: $\alpha = n \cdot p$;
ii) timpii între două apariţii succesive sunt variabile aleatorii independente, identic distribuite după o lege exponenţială de parametru: $\lambda = \alpha/\Delta T$.

- Aplicaţii:numărul emisiunilor din substanţele radioactive, numărul mesajelor sosite într-un sistem de comunicaţii, numărul cererilor de conexiuni într-o reţea de comunicaţie, numărul defectelor dintr-un dispozitiv electronic etc.

5.6.6. Distribuţia uniformă discretă

- Se presupune că: variabila aleatorie X ia valori într-un interval finit $\{1,2,\cdots,n\}$ de numere naturale. Toate valorile sunt echiprobabile.
- Spaţiul realizărilor: $S_X = \{1,2,\cdots,n\}$.
- Parametrii: n.
- Probabilităţi elementare: $p_X(k) = P[X = k] = 1/n$ pentru $k = \overline{1,n}$
- Media: $E[X] = (n+1)/2$;
- Varianţa: $VAR[X] = (n^2-1)/12$.
- Funcţia generatoare: $G_X(z) = \frac{1}{n} \cdot \frac{1-z^n}{1-z}$.
- Aplicaţii: evaluarea raportului semnal - zgomot introdus de procesul de cuantizare etc.

5.6.7. Distribuţia uniformă continuă

Se presupune că: variabila aleatorie X ia valori reale în intervalul $[a,b]$. Probabilitatea ca variabila aleatorie să ia o valoare într-un subinterval este proporţională cu lungimea intervalului.

- Spaţiul realizărilor: $S_X = [a,b]$.
- Parametrii: a şi b, $(a < b)$.
- Funcţia densitate de probabilitate: $f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{pentru } a \le x \le b \\ 0 & \text{în rest} \end{cases}$
- Funcţia de distribuţie: $F_X(x) = \begin{cases} 0 & \text{pentru } x < a \\ \dfrac{x-a}{b-a} & \text{pentru } a \le x \le b \\ 1 & \text{pentru } x > b \end{cases}$
- Media: $\mathrm{E}[X] = \dfrac{a+b}{2}$.
- Varianţa: $\mathrm{VAR}[X] = \dfrac{(b-a)^2}{12}$.
- Funcţia caracteristică: $\Phi_X(\omega) = \dfrac{e^{j\omega b} - e^{j\omega a}}{j\omega \cdot (b-a)}$.
- Aplicaţii: valori distribuite uniform în intervalul $[0,1]$ sunt utilizate pentru a genera numere ce corespund altor distribuţii, [Gar89], [Jer91].

5.6.8. Distribuţia exponenţială

- Spaţiul realizărilor: $S_X = [0,\infty)$.
- Parametrii: $\lambda > 0$.
- Funcţia densitate de probabilitate: $f_X(x) = \lambda \cdot \exp(-\lambda \cdot x)$.
- Funcţia de distribuţie: $F_X(x) = 1 - \exp(-\lambda \cdot x)$.
- Media: $\mathrm{E}[X] = 1/\lambda$;
- Varianţa: $\mathrm{VAR}[X] = 1/\lambda^2$.
- Funcţia caracteristică: $\Phi_X(\omega) = \lambda/(\lambda - j\omega)$.

Observaţii: i) reprezintă forma limită a distribuţiei geometrice;
ii) **este unica distribuţie continuă ce deţine proprietatea de lipsă a memorie**.

- Aplicaţii: lungimea mesajelor şi timpul dintre două apariţii succesive de evenimente (cereri de serviciu) în sistemele de comunicaţii; timpul de funcţionare (fiabilitatea sistemelor şi a componentelor).

5.6.9. Distribuţia Pareto

- Spaţiul realizărilor: $S_X = [k,+\infty)$.
- Parametri: parametrul de "formă" $\alpha > 0$ şi parametrul de "scară" $k > 0$
- Funcţia densitate de probabilitate: $f_X(x) = \alpha \cdot k^{\alpha} / x^{\alpha+1}$, cu $x \ge k$ (figura 5.6.1)

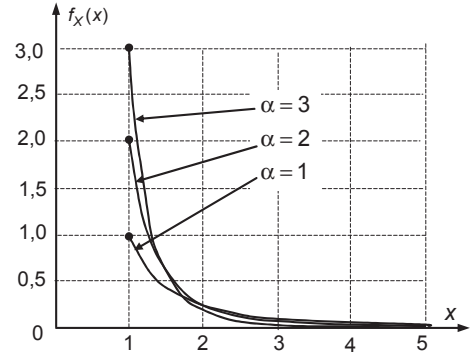


Figura 5.6.1: Densitatea de probabilitate a unei variabile aleatorii Pareto (k = 1)

- Funcţia de distribuţie: $F_X(x) = 1 - (k/x)^{\alpha}$.
- Media: $\mathrm{E}[X] = \begin{cases} \alpha \cdot k/(\alpha-1) & \text{pentru } \alpha > 1 \\ \infty & \text{pentru } \alpha \le 1 \end{cases}$;
- Varianţa: $\mathrm{VAR}[X] = \begin{cases} \dfrac{\alpha \cdot k^2}{(\alpha-1)^2 \cdot (\alpha-2)} & \text{cu } \alpha > 2 \\ \infty & \text{cu } \alpha \le 2 \end{cases}$
- Funcţia caracteristică: $\Phi_X(\omega) = \alpha \cdot (-ik \cdot t)^{\alpha} \cdot \Gamma(-\alpha, -ik \cdot t)$, unde $\Gamma(*,\circ)$ este ***funcţia gama incompletă***.

Observaţie: spre deosebire de distribuţia anterioară şi cea ulterioară, cu "*căderi exponenţiale*", valori semnificative ale densităţii de probabilitate vizează un interval sporit din spaţiu de eşantionare (***heavy tail distribution***)

- Aplicaţii: distribuţia lungimii fişierelor transferate prin Internet folosind protocolul TCP (***regula 80-20***: 20% din fişiere reprezintă 80% din trafic, adică câteva fişiere mari şi multe fişiere mici).

5.6.10. Distribuţia gaussiană (normală)

Se presupune că: o variabilă aleatorie X ce este determinată de o sumă în număr mare de cauze independente tinde să urmeze o distribuţie Gauss (conform teoremei limită centrală).

- Spaţiul realizărilor: $S_X = (-\infty, +\infty)$.
- Parametrii: μ şi σ.
- Funcţia densitate de probabilitate:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}. \text{ cu } \sigma > 0$$

- Funcţia de distribuţie:

$$F_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{x}\exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}dx \underset{t=\frac{x-\mu}{\sigma}}{=} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(x-\mu)/\sigma}\exp\left(\frac{-t^2}{2}\right)dt \overset{\text{notat.}}{=} \Phi\left(\frac{x-\mu}{\sigma}\right).$$

unde $\Phi(x)$ este funcţia de distribuţie a unei variabile aleatorii gaussiene de parametri $\mu = 0$ şi $\sigma = 1$.

În practică, se obişnuieşte a se lucra cu funcţia $Q(x)$ ("coada" distribuţiei) definită astfel:

$$Q(x) = 1 - \Phi(x) = 1/\sqrt{2\pi}\cdot\int_x^{\infty}\exp\left(-t^2/2\right)dt\,.$$

Valorile acestei funcţii se obţin prin integrare numerică. Urmare a simetriei pare a lui $f_X(x)$ (vezi figura 5.6.2), rezultă:

$$Q(x) = \Phi(-x) \quad \text{şi} \quad Q(-x) = 1 - \Phi(-x)$$

Pentru $0 < x < \infty$ se poate utiliza aproximaţia (cu $a = 1/\pi$ şi $b = 2\pi$) [Borj79]:

$$Q(x) \approx \left\{\left[(1-a)x + a\sqrt{x^2+b}\right]\sqrt{2\pi}\right\}^{-1}\cdot\exp\left(-x^2/2\right) \qquad (5.6.1)$$

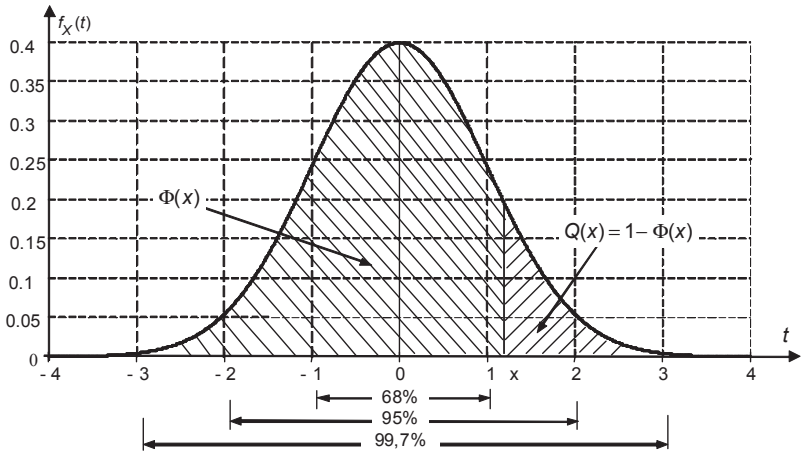


Figura 5.6.2: Funcţia de distribuţie a unei variabile aleatorii Gauss

Tabelul 5.6.1 prezintă câteva valori ale funcţiei $Q(x)$ obţinute prin ambele metode, iar tabelul 5.6.2 oferă valorile lui *x* pentru care $Q(x) = 10^{-k}$.

Tabelul 5.6.1: Comparaţie între *Q*(*x*) şi aproximarea dată de relaţia (5.6.1)

x	*Q*(*x*)	Aproximarea	*x*	*Q*(*x*)	Aproximarea
0,0	$5{,}00.10^{-1}$	$5{,}00.10^{-1}$	**2,7**	$3{,}47.10^{-3}$	$3{,}46.10^{-3}$
0,1	$4{,}60.10^{-1}$	$4{,}58.10^{-1}$	**2,8**	$2{,}56.10^{-3}$	$2{,}55.10^{-3}$
0,2	$4{,}21.10^{-1}$	$4{,}17.10^{-1}$	**2,9**	$1{,}87.10^{-3}$	$1{,}86.10^{-3}$
0,3	$3{,}82.10^{-1}$	$3{,}78.10^{-1}$	**3,0**	$1{,}35.10^{-3}$	$1{,}35.10^{-3}$
0,4	$3{,}45.10^{-1}$	$3{,}41.10^{-1}$	**3,1**	$9{,}68.10^{-4}$	$9{,}66.10^{-4}$
0,5	$3{,}09.10^{-1}$	$3{,}05.10^{-1}$	**3,2**	$6{,}87.10^{-4}$	$6{,}86.10^{-4}$
0,6	$2{,}74.10^{-1}$	$2{,}71.10^{-1}$	**3,3**	$4{,}83.10^{-4}$	$4{,}83.10^{-4}$
0,7	$2{,}42.10^{-1}$	$2{,}39.10^{-1}$	**3,4**	$3{,}37.10^{-4}$	$3{,}36.10^{-4}$
0,8	$2{,}12.10^{-1}$	$2{,}09.10^{-1}$	**3,5**	$2{,}33.10^{-4}$	$2{,}32.10^{-4}$
0,9	$1{,}84.10^{-1}$	$1{,}82.10^{-1}$	**3,6**	$1{,}59.10^{-4}$	$1{,}59.10^{-4}$
1,0	$1{,}59.10^{-1}$	$1{,}57.10^{-1}$	**3,7**	$1{,}08.10^{-4}$	$1{,}08.10^{-4}$
1,1	$1{,}36.10^{-1}$	$1{,}34.10^{-1}$	**3,8**	$7{,}24.10^{-5}$	$7{,}23.10^{-5}$
1,2	$1{,}15.10^{-1}$	$1{,}14.10^{-1}$	**3,9**	$4{,}81.10^{-5}$	$4{,}81.10^{-5}$
1,3	$9{,}68.10^{-2}$	$9{,}60.10^{-2}$	**4,0**	$3{,}17.10^{-5}$	$3{,}16.10^{-5}$
1,4	$8{,}08.10^{-2}$	$8{,}01.10^{-2}$	**4,5**	$3{,}40.10^{-6}$	$3{,}40.10^{-6}$
1,5	$6{,}68.10^{-2}$	$6{,}63.10^{-2}$	**5,0**	$2{,}87.10^{-7}$	$2{,}87.10^{-7}$
1,6	$5{,}48.10^{-2}$	$5{,}44.10^{-2}$	**5,5**	$1{,}90.10^{-8}$	$1{,}90.10^{-8}$
1,7	$4{,}46.10^{-2}$	$4{,}43.10^{-2}$	**6,0**	$9{,}87.10^{-10}$	$9{,}86.10^{-10}$
1,8	$3{,}59.10^{-2}$	$3{,}57.10^{-2}$	**6,5**	$4{,}02.10^{-11}$	$4{,}02.10^{-11}$
1,9	$2{,}87.10^{-2}$	$2{,}86.10^{-2}$	**7,0**	$1{,}28.10^{-12}$	$1{,}28.10^{-12}$
2,0	$2{,}28.10^{-2}$	$2{,}26.10^{-2}$	**7,5**	$3{,}19.10^{-14}$	$3{,}19.10^{-14}$
2,1	$1{,}79.10^{-2}$	$1{,}78.10^{-2}$	**8,0**	$6{,}22.10^{-16}$	$6{,}22.10^{-16}$
2,2	$1{,}39.10^{-2}$	$1{,}39.10^{-2}$	**8,5**	$9{,}48.10^{-18}$	$9{,}48.10^{-18}$
2,3	$1{,}07.10^{-2}$	$1{,}07.10^{-2}$	**9,0**	$1{,}13.10^{-19}$	$1{,}13.10^{-19}$
2,4	$8{,}20.10^{-3}$	$8{,}17.10^{-3}$	**9,5**	$1{,}05.10^{-21}$	$1{,}05.10^{-21}$
2,5	$6{,}21.10^{-3}$	$6{,}19.10^{-3}$	**10,0**	$7{,}62.10^{-24}$	$7{,}62.10^{-24}$
2,6	$4{,}66.10^{-3}$	$4{,}65.10^{-3}$			

Tabelul 5.6.2: Valorile argumentului *x* al funcţiei $Q(x) = 10^{-k}$ pentru câteva valori ale exponentului *k*

k	$x = Q^{-1}(10^{-k})$	*k*	$x = Q^{-1}(10^{-k})$
1	1.2815	**6**	4.7535
2	2.3263	**7**	5.1993
3	3.0902	**8**	5.6120
4	3.7190	**9**	5.9978
5	4.2649	**10**	6.3613

- Media: $\mathrm{E}[X] = \mu$; Varianţa: $\mathrm{VAR}[X] = \sigma^2$.
- Funcţia caracteristică: $\Phi_X(\omega) = \exp\left(j\mu\omega - \sigma^2\omega^2/2\right)$

- Aplicaţii: - modelează distribuţia de amplitudine a zgomotului termic;
  - modelează funcţionarea multiplexoarelor statistice;
  - aproximează distribuţia binomială etc.

5.6.11. Distribuţia Gama

- Spaţiul realizărilor: $S_X = (0,\infty)$.
- Parametrii: $\alpha > 0$ şi $\lambda > 0$
- Funcţia densitate de probabilitate : $f_X(x) = \lambda(\lambda \cdot x)^{\alpha-1} \cdot \dfrac{\exp(\lambda \cdot x)}{\Gamma(\alpha)}$

unde $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} \cdot e^{-y} dy$, cu proprietăţile $\begin{cases} \Gamma(1/2) = \sqrt{\pi} \\ \Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha), \text{ pentru } \alpha \in N \\ \Gamma(\alpha) = (\alpha-1)! \end{cases}$.

- Media: $E[X] = \alpha/\lambda$;
- Varianţa: $VAR[X] = \alpha/\lambda^2$.
- Funcţia caracteristică: $\Phi_X(\omega) = (1 - j\omega/\lambda)^{-\alpha}$.
- Aplicaţii: distribuţia gama, prin intermediul parametrilor α şi λ, poate genera o varietate de curbe (vezi figura 5.6.3) prin care se pot modela diverse experimente.

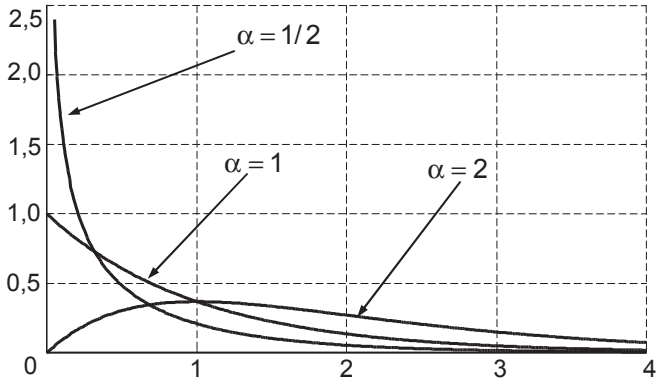


Figura 5.6.3: Distribuţia Gama: funcţia densitate de probabilitate

O serie de distribuţii sunt cazuri particulare ale distribuţiei gama. Astfel:
- pentru $\alpha = 1$ se obţine distribuţia exponenţială;
- pentru $\lambda = 1/2$ şi $\alpha = k/2$ cu $k \in N^*$ se obţine distribuţia χ^2;
- pentru $\alpha = m$, $m \in N^*$ se obţine distribuţia *m-Erlang.*

5.6.12. Distribuţia m-Erlang

Caz particular al distribuţiei Gama: $\alpha = m$, $m \in N^*$.

- Spaţiul realizărilor: $S_X = (0,\infty)$.
- Parametrii: $m \in N^*$ şi $\lambda > 0$.
- Funcţia densitate de probabilitate: $f_X(x) = \dfrac{\lambda \cdot (\lambda \cdot x)^{m-1}}{(m-1)!} \cdot \exp(-\lambda \cdot x)$.
- Funcţia de distribuţie: $F_X(x) = 1 - \sum_{k=0}^{m-1} \dfrac{(\lambda \cdot x)^k}{k!} \exp(-\lambda \cdot x)$.
- Media: $E[X] = \dfrac{m}{\lambda}$.
- Varianţa: $VAR[X] = \dfrac{m}{\lambda^2}$
- Funcţia caracteristică: $\Phi_X(\omega) = (1 - j\omega/\lambda)^{-m}$.
- Aplicaţii: O variabilă aleatorie m-Erlang se obţine adunând m variabile aleatorii exponenţiale independente, identic distribuite, de parametru λ. De aceea, ea este un model general al timpilor de aşteptare în sistemele cu aşteptare, al timpilor de viaţă în studiile de fiabilitate etc.

5.6.13. Distribuţia χ^2

Caz particular al distribuţiei Gama: $\lambda = 1/2$, $\alpha = m/2$ şi $m \in N^*$.

- Spaţiul realizărilor: $S_X = (0,\infty)$.
- Parametrii: m (grade de libertate).
- Funcţia densitate de probabilitate: $f_X(x) = \dfrac{x^{(m-2)/2} \cdot e^{-x/2}}{2^{m/2} \cdot \Gamma(m/2)}$.
- Funcţia de distribuţie: $F_X(x)$, vezi tabelul 5.6.3.
- Media: $E[X] = m$; Varianţa: $VAR[X] = 2m$.
- Funcţia caracteristică: $\Phi_X(\omega) = \dfrac{\lambda}{(1 - j \cdot 2\omega)^{m/2}}$.

Observaţie: $X = \sum_{i=1}^{m} X_i^2$, (cu X_i sunt variabile aleatorii Gauss, i.i.d. cu $\mu_i = 0$ şi $\sigma_i = 1$) este o variabilă aleatorie χ^2 cu m grade de libertate.

- Aplicaţii: Se utilizează în analiza varianţei estimatorilor obţinuţi pe baza măsurătorilor de putere şi în aprecierea aproximării datelor experimentale de către modelul probabilistic presupus etc.

Tabelul 5.6.3: Valorile argumentului funcţiei de distribuţie χ^2 pentru nivele specifice ale acesteia şi diverse grade de libertate

	$F(x)$						
	1%	**5%**	**25%**	**50%**	**75%**	**95%**	**99%**
***m* = 1**	0.00016	0.00393	0.1015	0.4549	1.323	3.841	6.635
***m* = 2**	0.02010	0.1026	0.5753	1.386	2.773	5.991	9.210
***m* = 3**	0.1148	0.3518	1.213	2.366	4.108	7.815	11.34
***m* = 4**	0.2971	0.7107	1.923	3.357	5.385	9.488	13.28
***m* = 5**	0.5543	1.1455	2.675	4.351	6.626	11.07	15.09
***m* = 6**	0.8720	1.635	3.455	5.348	7.841	12.59	16.81
***m* = 7**	1.239	2.167	4.255	6.346	9.037	14.07	18.48
***m* = 8**	1.646	2.733	5.071	7.344	10.22	15.51	20.09
***m* = 9**	2.088	3.325	5.899	8.343	11.39	16.92	21.67
***m* = 10**	2.558	3.940	6.737	9.342	12.55	18.31	23.21
***m* = 11**	3.053	4.575	7.584	10.34	13.70	19.68	24.73
***m* = 12**	3.571	5.226	8.438	11.34	14.84	21.03	26.22
***m* = 15**	5.229	7.261	11.04	14.34	18.25	25.00	30.58
***m* = 20**	8.260	10.85	15.45	19.34	23.83	31.41	37.57
***m* = 30**	14.95	18.49	24.48	29.34	34.80	43.77	50.89
***m* = 50**	29.71	34.76	42.94	49.33	56.33	67.50	76.15
***m* > 30**	$m+\sqrt{2m}\cdot x+2/3(x^2-1)+o(1/\sqrt{m})$						
***x* =**	– 5.33	– 1.64	– 0.675	0.00	0.675	1.64	2.33

Tabelul 5.6.4: Valorile argumentului *z* pentru probabilităţile specificate, $P[-z \le X \le z]$, ale unei variabile aleatorii t, cu diverse grade de libertate

***m* – 1**	$P[-z \le X \le z]$		
	0.90	**0.95**	**0.99**
1	6.314	12.706	63.657
2	2.920	4.303	9.925
3	2.353	3.182	5.841
4	2.132	2.776	4.604
5	2.015	2.571	4.032
6	1.943	2.447	3.707
7	1.895	2.365	3.499
8	1.860	2.306	3.355
9	1.833	2.262	3.250
10	1.812	2.228	3.169
15	1.753	2.131	2.947
20	1.725	2.086	2.845
30	1.697	2.042	2.750
40	1.684	2.021	2.704
60	1.671	2.000	2.660
∞	1.645	1.960	2.576

5.6.14. Distribuţia Student, t

Se presupune că:

i) $X=\dfrac{Z}{\sqrt{\dfrac{Y_m}{m}}}$, unde Z este o variabilă aleatorie gaussiană cu $\mu_Z=0$ şi $\sigma_Z=1$, iar Y_m este o variabilă aleatorie χ^2 cu m grade de libertate.

ii) Z şi Y_m sunt independente.

- Spaţiul realizărilor: $S_X=(-\infty,+\infty)$.
- Parametrii: m.
- Funcţia de distribuţie: $F_X(x)$ vezi tabelul 5.6.4.
- Media: $E[X]=0$; Varianţa: $VAR[X]=m/(m-2)$, dacă $m>0$
- Funcţia densitate de probabilitate: $f_X(x)=\dfrac{\Gamma((m+1)/2)}{\sqrt{m\pi}\cdot\Gamma(m/2)\cdot\sqrt{\left(1+x^2/m\right)^{m+1}}}$
- Aplicaţii: Stabileşte intervale de încredere pentru estimarea mediei când varianţa nu este cunoscută, sau când eşantioanele conţin un număr mic de rezultate, sub 20 etc.

5.6.15. Distribuţia Fisher - Snedecor (F)

Se presupune că U şi V sunt două variabile aleatorii χ^2 cu m_1, respectiv m_2 grade de libertate. Variabila aleatorie F se obţine c: $F=\dfrac{U/m_1}{V/m_2}$.

- Spaţiul realizărilor: $S_F=(0,+\infty)$.
- Parametri: m_1 şi m_2. Funcţia densitate de probabilitate:

$$f_F(x)=\frac{(m_1/m_2)^{m_1/2}\cdot x^{m_2/2-1}}{\Gamma\left(\frac{m_1}{2}\right)\cdot\Gamma\left(\frac{m_2}{2}\right)\cdot\left(1+\frac{m_1}{m_2}\cdot x\right)^{(m_1+m_2)/2}}\cdot\Gamma\left(\frac{m_1+m_2}{2}\right)$$

- Funcţia de distribuţie: $F_F(x)$ tabelată, vezi [Andr95].
- Media: $E[F]=\dfrac{m_2}{m_2-2}$, dacă $m_2>2$.
- Varianţa: $VAR[F]=\dfrac{2m_2^2\cdot(m_1+m_2-2)}{m_1\cdot(m_2-2)^2\cdot(m_2-4)}$, dacă $m_2>4$.
- Aplicaţii: Se utilizează în analiza proprietăţilor estimatorilor varianţei şi covarianţei, a nivelelor de putere, a raportului semnal-zgomot etc.

5.6.16. Distribuţia Laplace

Spaţiul realizărilor: $S_X = (-\infty, +\infty)$.

- Parametrii: $\alpha > 0$.
- Funcţia densitate de probabilitate: $f_X(x) = \frac{\alpha}{2} \cdot e^{-\alpha|x|}$.
- Funcţia de distribuţie: $F_X(x) = \begin{cases} 1/2 \cdot e^{\alpha x} & \text{pentru } x < 0 \\ 1 - 1/2 \cdot e^{-\alpha x} & \text{pentru } x \geq 0 \end{cases}$
- Media: $E[X] = 0$. Varianţa: $VAR[X] = 2/\alpha^2$.
- Funcţia caracteristică: $\Phi_X(\omega) = \frac{\alpha^2}{\omega^2 + \alpha^2}$.

5.6.17. Distribuţia Rayleigh

Dacă X_1 şi X_2 sunt variabile aleatorii Gauss independente, identic distribuite de medie 0 şi varianţă σ^2, atunci variabila aleatorie X (Rayleigh) se obţine cu relaţia: $X = \sqrt{X_1^2 + X_2^2}$.

- Spaţiul realizărilor: $S_X = [0, \infty)$.
- Parametrii: σ^2.
- Funcţia densitate de probabilitate : $f_X(x) = \frac{x}{\sigma^2} \cdot \exp\left(\frac{-x^2}{2\sigma^2}\right)$.
- Funcţia de distribuţie: $F_X(x) = 1 - \exp\left(\frac{-x^2}{2\sigma^2}\right)$.
- Media: $E[X] = \sigma \cdot \sqrt{\pi/2}$; Varianţa: $VAR[X] = \sigma^2 \cdot (2 - \pi/2)$.
- Aplicaţii: fadingul în canalul de comunicaţie, anvelopa zgomotului gaussian de bandă îngustă.

5.6.18. Distribuţia Cauchy

Nu se fac presupuneri.

- Spaţiul realizărilor: $S_X = (-\infty, +\infty)$.
- Parametrii: $\alpha > 0$.
- Funcţia densitate de probabilitate : $f_X(x) = \frac{\alpha/\pi}{x^2 + \alpha^2}$.
- $E[X]$ şi $VAR[X]$ nu există.
- Funcţia caracteristică: $\Phi_X(x) = \exp(-\alpha \cdot |\omega|)$.

5.7 Metode de aproximare

În multe aplicaţii, ne putem găsi în următoarele situaţii:

- variabila aleatorie urmărită nu este complet specificată, ci sunt cunoscute doar media, varianţa şi eventual câteva momente centrate de ordin superior, $E\left[(X - \mu_X)^k\right]$, pentru $k > 2$;
- funcţia densitate de probabilitate este cunoscută, dar este imposibil de integrat (vezi distribuţia Gauss);
- variabila aleatorie urmărită este o sumă ponderată de variabile aleatorii (cauze) independente, mai mult sau mai puţin specificate.

În aceste cazuri, pentru a deţine informaţii relative la comportamentul variabilei aleatorii, se utilizează diferite tehnici de aproximare. Pe lângă metodele prezentate în cele ce urmează, subcapitolele 3.3.5-3.3.7 oferă, la rândul lor, alte posibilităţi de aproximare.

5.7.1. Inegalitatea Markov

Se presupune că variabila aleatorie X este pozitivă şi se cunoaşte media $E[X]$. În aceste condiţii. inegalitatea Markov se enunţă în modul următor:

$$P[X \geq a] \leq E[X]/a \qquad (5.7.1)$$

Expresia ce urmează este precizată ca demonstraţie a inegalităţii:

$$E[X] = \int_0^\infty x \cdot f_X(x)\,dx = \int_0^a \cdot f_X(x)\,dx + \int_a^\infty x \cdot f_X(x)\,dx$$
$$\geq \int_a^\infty t \cdot f_X(x)\,dx \geq \int_a^\infty a \cdot f_X(x)\,dx = a \cdot P[X \geq a]$$

Aplicaţia 5.7.1

Media înălţimii elevilor într-o clasă este 1,50 metri. Determinaţi marginea superioară a probabilităţii ca un elev să fie mai înalt de 2 metri.

Rezolvare: $P[X \geq 2] \leq 1{,}5/2 = 0{,}75$.

Observaţie: Rezultatul obţinut îl considerăm lipsit de sens deoarece în cazul acestei variabile aleatorii (X = înălţimea copiilor) deţinem şi informaţii relative la varianţa sa (variabilitatea biologică a înălţimii copiilor faţă de medie), pe care metoda nu le foloseşte.

5.7.2. Inegalitatea Cebâşev

Se presupune că se cunosc valorile mediei $E[X] = m$ şi varianţei $VAR[X] = \sigma^2$ a unei variabile X, iar inegalitatea Cebâşev se enunţă în modul următor:

$$P[|X - m| \geq a] \leq \sigma^2/a^2 \qquad (5.7.2)$$

Demonstraţie: pentru variabila aleatorie X, următoarele două evenimente sunt echivalente: $\{|X| \geq a\}$ şi $\{X^2 \geq a^2\}$. Aplicăm inegalitatea Markov pentru variabila aleatorie X^2 şi obţinem: $P[X^2 \geq a^2] \leq \mathrm{E}[X^2]/a^2$ sau $P[|X| \geq a] \leq \mathrm{E}[X^2]/a^2$. Se face substituţia $X \leftarrow X - m$ şi din ultima expresie rezultă că: $P[|X-m| \geq a] \leq \mathrm{E}[(X-m)^2]/a^2 = \sigma^2/a^2$, q. e. d.

Observaţie: Pentru o variabilă aleatorie cu $\sigma = 0$, $P[|X-m| \geq a] = 0$, pentru orice a, deci $P[X = m] = 1$, adică, în aproape toate experimentele, variabila aleatorie ia valoarea m.

Aplicaţia 5.7.2

Media şi deviaţia standard a timpului de răspuns al unui calculator multiserver este m sec, respectiv σ sec. Estimaţi, utilizând inegalitatea Cebâşev, probabilitatea ca timpul de răspuns să se abată cu $k \cdot \sigma$ sec faţă de medie.

Comparaţi marginea superioară rezultată cu valoarea exactă a probabilităţii în cazul distribuţiei Gauss. Caz particular $k = 2$.

Rezolvare: $$P[|X-m| > k \cdot \sigma] \leq \frac{\sigma^2}{k^2 \cdot \sigma^2} = \frac{1}{k^2} = \frac{1}{4} = 0{,}25$$

$$P[|X-m| > k\sigma] \leq 2 \cdot Q(k) = 0{,}0456 << 0{,}25$$

Observaţie: şi această inegalitate, în general, este mult acoperitoare.

5.7.3. Inegalitatea Chernoff

Se presupune că se cunoaşte $\mathrm{E}[\exp(t \cdot X)]$.

Enunţ:

$$P[X \geq a] \leq \min_{t \geq 0}\left(\mathrm{e}^{-t \cdot a}\right) \cdot \mathrm{E}\left[\mathrm{e}^{t \cdot X}\right] \leq \min_{t \geq 0}\left\{\exp\left(-t \cdot a + \ln \mathrm{E}\left[\mathrm{e}^{t \cdot X}\right]\right)\right\} \quad (5.7.3)$$

Demonstraţie:

Se defineşte o variabilă aleatorie $Y = \begin{cases} 1 & \text{pentru } X \geq a \\ 0 & \text{pentru } X < a \end{cases}$. În acest caz, pentru $\forall t \geq 0$, avem: $\exp(t \cdot X) \geq \exp(t \cdot a) \cdot Y$. Mediind obţinem că:

$$\mathrm{E}[\exp(t \cdot X)] \geq \exp(t) \cdot \mathrm{E}[Y],$$

cu: $\mathrm{E}[Y] = 1 \cdot P[Y=1] + 0 \cdot P[Y=0] = 1 \cdot P[X \geq a]$, evenimentele: $\{Y = 1\}$ şi $\{X \geq a\}$ fiind echivalente.

Deci: $\mathrm{E}[\exp(t \cdot X)] \geq \exp(t \cdot a) \cdot P[X \geq a]$ şi această inegalitate fiind valabilă pentru $\forall t$, rezultă că :

$$P[X \geq a] \leq \min_{t \geq 0}\{\exp(-t \cdot a)\} \cdot \mathrm{E}[\exp(t \cdot X)] \quad \text{q.e.d.}$$

capitolul 6

Variabile aleatorii vectoriale

Variabilă aleatorie vectorială, $\mathbf{X} = (X_1, X_2, \cdots, X_n)$, este o funcţie care alocă un vector de n numere reale fiecărei realizări ζ din spaţiul de eşantionare, S, al unui experiment aleator. Elementele vectorului pot reprezenta observaţii măsurabile, ce caracterizează o realizare, câteva măsurători repetate ale unei anumite mărimi etc.

Aplicaţia 6.1

Daţi exemple de experimente care generează variabile aleatorii vectoriale.

Rezolvare: Fie eşantionarea cu perioadă T, a unui semnal $X(t)$. Eşantionul, $X_k = X(k \cdot T)$, luat la momentul $k \cdot T$ reprezintă o variabilă aleatorie, iar vectorul alcătuit din primele *n* eşantioane, adică $\mathbf{X} = (X_1, X_2, \cdots, X_n)$, este o *variabilă aleatorie vectorială*.

Observaţie: Noţiunea de *eşantion* se întâlneşte atât în statistică, cât şi în domeniul prelucrării semnalelor electrice. Semnificaţiile sunt însă puţin diferite, în primul caz reprezentând un ansamblu de realizări ale unui fenomen aleatoriu, iar în al doilea caz constituind valoarea unei mărimi electrice măsurate la un moment dat.

* * *

O **variabilă aleatorie vectorială** poate fi **continuă,** dacă toate componentele sale sunt variabile aleatorii continue, **discretă**, dacă toate componentele sale sunt discrete, sau **mixtă**, dacă o parte din componente sunt variabile aleatorii continue şi restul discrete.

Evenimentelor ce implică o variabilă aleatorie vectorială cu spaţiu de eşantionare *n*-dimensional, le corespund regiuni în acest spaţiu S.

Aplicaţia 6.2

Fie variabila aleatorie vectorială (bidimensională) $\mathbf{V}=(X,Y)$. Presupunând că $\mathrm{S_V}=\square\times\square$, să se determine regiunile ce corespund evenimentelor:

i) $A=\{X+Y\le 10\}$; ii) $B=\{\min(X,Y)\le 5\}$; iii) $C=\{X^2+Y^2\le 100\}$.

Rezolvare: regiunile de interes sunt precizate în figura 6.1.

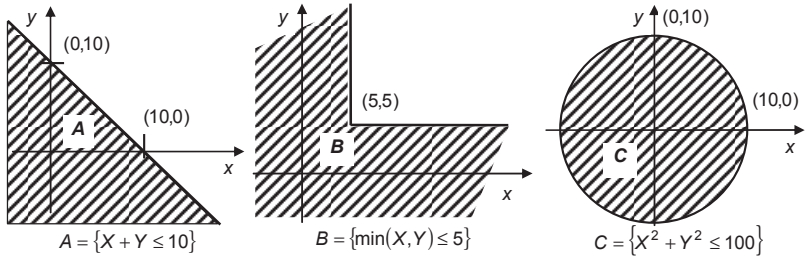


Figura 6.1: Exemple de evenimente bidimensionale

* * *

În analiza variabilelor aleatorii vectoriale prezintă interes evenimentele (notate generic cu A) ce au **forma produs**:

$$A=\{X_1\in A_1\}\&\{X_2\in A_2\}\&\ldots\&\{X_n\in A_n\} \qquad (6.1)$$

unde $A_k=$ eveniment unidimensional ce implică doar variabila aleatorie X_k.

Pentru simplificarea scrierii expresiilor unor astfel de evenimente, se poate folosi şi forma:

$$A=X_1\in A_1\ \&\ X_2\in A_2\ \&\cdots\&\ X_n\in A_n \qquad (6.2)$$

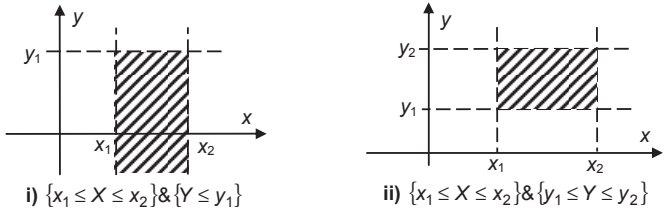


Figura 6.2: Exemple de evenimente bidimensionale cu formă produs

Figura 6.2 oferă două exemple de evenimente bidimensionale ce sunt exprimabile în formă produs.

Evenimentele ce nu au forma produs pot fi aproximate, la limită identificate, cu reuniuni de evenimente produs (vezi figura 6.3). Astfel, dacă B este un eveniment fără formă produs şi $B_1,B_2,\cdots$ evenimente cu formă produs, disjuncte, a căror reuniune îl aproximează, atunci probabilitatea evenimentului *B* este dată de relaţia:

$$P[B]\cong P\left[\bigcup\nolimits_k B_k\right]=\sum\nolimits_k P[B_k] \qquad (6.3)$$

La limită, când evenimentele B_k devin arbitrar de "fine", relaţia devine exactă.

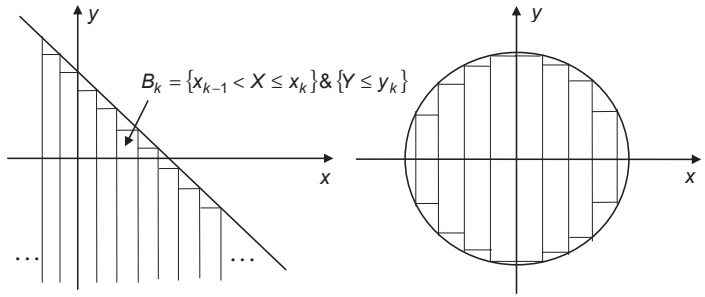


Figura 6.3: Exemple de evenimente fără formă produs şi aproximarea lor cu reuniuni de evenimente disjuncte cu formă produs

Aplicaţia 6.3

Scrieţi sub formă produs evenimentele din exemplul anterior (*Aplicaţia 6.2*).

Rezolvare: $B=\{\min(X,Y)\le 5\}=\{X\le 5\ \&\ Y<\infty\}\cup\{X>5\ \&\ Y\le 5\}$

* * *

Variabilele aleatorii $X_1,X_2,\cdots,X_n$ ca elemente ale vectorului $\mathbf{X}$, sunt **independente** dacă:

$$P[X_1\in A_1,X_2\in A_2,\ldots,X_n\in A_n]=P[X_1\in A_1]\cdot P[X_2\in A_2]\cdot\ldots\cdot P[X_n\in A_n] \quad (6.4)$$

deci dacă fiecare variabilă aleatorie X_k este implicată doar de evenimentul unidimensional A_k definit pe dimensiunea respectivă.

6.1 Variabile aleatorii bidimensionale

Înainte de începerea analizei acestui caz, este important să se facă observaţia că procedeele ce vor fi utilizate în continuare, pentru a deduce principalele mărimi care descriu comportamentul aleatoriu al unei variabile

aleatorie bidimensionale, se pot extinde pentru variabile aleatorii vectoriale de orice dimensiune.

O variabilă aleatorie bidimensională este un vector ale cărui componente pot avea, în mod distinct, spaţiul de eşantionare discret sau continuu.

O variabilă aleatorie bidimensională discretă, $\mathbf{V}=(X,Y)$, adică cu ambele componente având un spaţiu de eşantionare discret, este complet descrisă de **probabilităţile elementare de legătură (simultane**):

$$p_{\mathbf{V}}(x_i,y_j)=P\left[X=x_i,Y=y_j\right] \text{ pentru } i,j=1,2,\cdots \qquad (6.1.1)$$

unde perechea (x_i,y_j) este un element al mulţimii de eşantionare $\mathrm{S}_{X,Y}$.

Probabilitatea de apariţie a oricărui eveniment, $A\subset \mathrm{S}_{X,Y}$, este:

$$P[A]=\sum_{(x_i,y_j)\in A} p(x_i,y_j)$$

deci: $$F_{X,Y}(x,y)=P[X\le x,Y\le y]=\sum_{\{(x_i,y_j):\,x_i\le x,\,y_j\le y\}} p_{X,Y}(x_i,y_j) \qquad (6.1.2)$$

Cunoscând funcţia de probabilitate a variabilei aleatorie bidimensionale discrete, $\mathbf{V}=(X,Y)$, apariţia evenimentelor elementare, ce implică doar o dimensiune (x sau y) este descrisă de **probabilităţile elementare marginale**:

$$p_X(x_j)=P\left[\left\{X=x_j,Y=y_1\right\}\cup\left\{X=x_j,Y=y_2\right\}\cup\cdots\right]=\sum_{k=1}^{+\infty}p(x_j,y_k) \qquad (6.1.3a)$$

respectiv: $$p_Y(y_j)=\sum_{k=1}^{+\infty}p(x_k,y_j) \qquad (6.1.3b)$$

Probabilitatea apariţiei unui eveniment unidimensional, B, definit pe această dimensiune este:

$$P[B]=\sum_{x_i\in B} p_X(x_i) \qquad (6.1.3c)$$

Observaţie: Cunoaşterea funcţiei de probabilitate a unei variabile aleatorie (X,Y) permite determinarea funcţiilor de probabilitate ale variabilelor aleatorii izolate, X şi Y. Reciproca însă, în general, nu este adevărată.

Aplicaţia 6.1.1

Numărul de octeţi, N, ai unui mesaj urmează o distribuţie geometrică de parametru p. Mesajul este segmentat în pachete de lungime maximă M octeţi. Fie Q, numărul de pachete complete şi R, numărul de octeţi din ultimul pachet.

Determinaţi funcţia de probabilitate a variabilei aleatorie $\mathbf{V}=(Q,R)$ precum şi probabilităţile marginale. Verificaţi dacă expresiile deduse verifică condiţia de normare.

Rezolvare: variabilele aleatorii *izolate* au următoarele spaţii de eşantionare:

- $\mathrm{S}_Q=\{0,1,\cdots\}$ şi $\mathrm{S}_R=\{0,1,\cdots,M-1\}$

- $p_V(q,r)=P[Q=q,R=r]=P[N=M\cdot q+r]\ne(1-p)\cdot p^{M\cdot q+r}$, $q\in\mathrm{N}$
- $$\begin{aligned}P[Q=q]&=P\left[N\in\{Mq,Mq+1,\cdots,Mq+(M-1)\}\right]\\&=P[N=Mq\vee N=Mq+1\vee\cdots\vee N=Mq+(M-1)]\\&=\sum_{k=0}^{M-1}(1-p)\,p^{Mq+k}=(1-p)\,p^{Mq}\sum_{k=0}^{M-1}p^k=(1-p^M)\left(p^M\right)^q,\ q\in\mathrm{N}\end{aligned}$$

Observaţii:

- funcţia de distribuţie marginală a variabilei aleatorie Q este o progresie geometrică de parametru p^M.

$$\begin{aligned}P[R=r]&=P\left[N\in\{r,M+r,2M+r,\cdots\}\right]\\&=\sum_{q=0}^{\infty}(1-p)\cdot p^{Mq+r}\cdot\frac{1-p}{1-p^M}\cdot p^r \text{ pentru } r=\overline{0,M-1}\end{aligned} \qquad (6.1.4)$$

- variabila aleatorie R, observată izolat, urmează o **distribuţie geometrică trunchiată**.

* * *

O variabilă aleatorie bidimensională continuă, $\mathbf{V}=(X,Y)$, este complet descrisă de **funcţia de distribuţie de legătură** (**simultană**) asociată ei:

$$F_{\mathbf{V}}(x_1,y_1)=F_{X,Y}(x_1,y_1)=P[X\le x_1,Y\le y_1] \qquad (6.1.5)$$

Ea reprezintă probabilitatea dreptunghiului (ca eveniment) semi-infinit, definit de punctul (x_1,y_1), ca în figura 6.1.1.

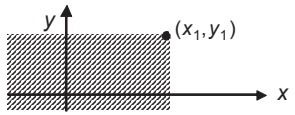


Figura 6.1.1: Evenimentul $\{X\le x_1,Y\le y_1\}$

Funcţia de distribuţie a unei variabile aleatorie vectoriale, (X,Y), are următoarele proprietăţi:

- $F_{X,Y}(x_1,y_1)\le F_{X,Y}(x_2,y_2)$, dacă $x_1\le x_2$ şi $y_1\le y_2$ (funcţia este monoton crescătoare pe direcţia "nord-est");
- $F_{X,Y}(-\infty,y)=F_{X,Y}(x,-\infty)=0$ (este imposibil ca variabilele aleatorii X sau Y să ia o valoare mai mică decât $-\infty$);
- $F_{X,Y}(\infty,\infty)=1$ (este sigur că X şi Y iau valori mai mici ca $+\infty$);
- $F_X(x)=F_{X,Y}(x,\infty)$ şi $F_Y(y)=F_{X,Y}(\infty,y)$ (reprezintă **funcţiile de distribuţie marginală** ale variabilelor aleatorii X, respectiv Y);

- $\lim_{x\to a^+} F_{X,Y}(x,y) = F_{X,Y}(a,y)$ şi $\lim_{y\to b^+} F_{X,Y}(x,y) = F_{X,Y}(x,b)$

(reprezintă continuitatea "de la nord" şi "de la est"; este indusă de continuitatea la dreapta a unei variabile aleatorii unidimensionale);

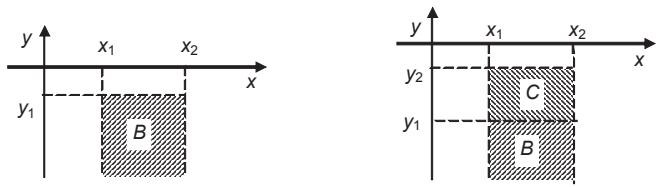
- probabilităţile evenimentelor ce implică o variabilă aleatorie bidimensională continuă se determină exprimând evenimentele ca reuniuni şi intersecţii de dreptunghiuri semi-infinite:

$$P[x_1 < x \le x_2 \,\&\, y_1 < y \le y_2] = F_{X,Y}(x_2,y_2) - F_{X,Y}(x_2,y_1) - F_{X,Y}(x_1,y_2 + F_{X,Y}(x_1,y_1)$$

Aplicaţia 6.1.2

Determinaţi probabilităţile evenimentelor:

$A = \{X \le x_1, Y \le y_1\}$, $B = \{x_1 < X \le x_2, Y \le y_1\}$ şi $C = \{x_1 < X \le x_2, y_1 < Y \le y_2\}$ pentru cazul în care comportamentul variabilei aleatorie (X,Y) este descris de funcţia de distribuţie $F_{X,Y}(x,y)$.



a. Evenimentul $\{x_1 < X \le x_2, Y \le y_1\}$ **b. Reuniune de evenimente disjuncte**

Figura 6.1.2

Rezolvare: i) $P[A] = P[X \le x_1, Y \le y_1] = F_{X,Y}(x_1,y_1)$

ii) Regiunea ce corespunde evenimentului B este reprezentat în figura 6.1.2a. Conform acesteia, există următoarea relaţie de echivalenţă între evenimente: $\{X \le x_2, Y \le y_1\} = \{X \le x_1, Y \le y_1\} \cup B$. Deoarece:

$\{X \le x_1, Y \le y_1\} \cap B = \varnothing$, rezultă că: $F_{X,Y}(x_2,y_1) = F_{X,Y}(x_1,y_1) + P[B]$,

şi deci: $P[x_1 < X \le x_2, Y \le y_1] = F_{X,Y}(x_2,y_1) - F_{X,Y}(x_1,y_1)$.

iii) Regiunea ce corespunde evenimentului C este reprezentat în figura 6.1.2b. Există următoarea relaţie de echivalenţă:

$$\{X \le x_2, Y \le y_2\} \equiv \{X \le x_1, Y \le y_2\} \cup B \cup C$$

în care evenimentele din membrul drept al egalităţii sunt disjuncte între ele, deci:

$$F_{X,Y}(x_2,y_2) = F_{X,Y}(x_1,y_2) + P[B] + P[C]$$

$$P[x_1 < X \le x_2 \,\&\, y_1 < Y \le y_2] = F_{X,Y}(x_2,y_2) - F_{X,Y}(x_1,y_2) - F_{X,Y}(x_2,y_1) + F_{X,Y}(x_1,y_1)$$

* * *

Aplicaţia 6.1.3

Funcţia de distribuţie de legătură a variabilei aleatorie (X,Y) este dată de expresia:

$$F_{X,Y}(x,y) = \begin{cases} (1-e^{-\alpha\cdot x})\cdot(1-e^{-\beta\cdot y}) & \text{pentru } x \ge 0 \text{ şi } y \ge 0 \\ 0 & \text{în rest} \end{cases}$$

Determinaţi funcţiile de distribuţie marginale.

Rezolvare:

$F_X(x) = F_{X,Y}(x,\infty) = \lim_{y\to\infty}(1-e^{-\alpha\cdot x})\cdot(1-e^{-\beta\cdot y}) = 1-e^{-\alpha\cdot x}$, pentru $x \ge 0$;

$F_Y(y) = 1-e^{-\beta\cdot y}$, pentru $y \ge 0$.

* * *

Funcţia densitate de probabilitate de legătură (simultană) a unei variabile aleatorie vectoriale bidimensionale, *continue,* $\mathbf{V} = (X,Y)$**,** *dacă există*, este:

$$f_{X,Y}(x,y) = \frac{\partial F_{X,Y}(x,y)}{\partial x\,\partial y} \qquad (6.1.6)$$

Restricţiile sunt legate de imposibilitatea de a deriva parţial funcţia de distribuţie (ea însăşi este discontinuă, derivata parţială după o variabilă este discontinuă etc.).

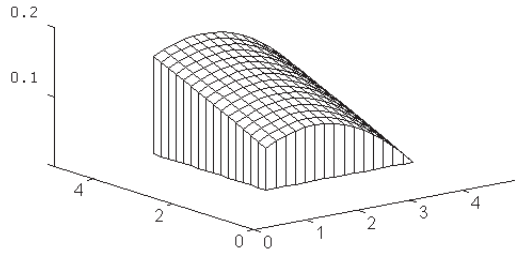


Figura 6.1.3: **Exemplu de funcţie de densitate de probabilitate a unei variabile aleatorie bidimensionale**

Exprimarea probabilităţii de apariţie a unui eveniment A prin intermediul densităţii de probabilitate este (vezi şi figura 6.1.6.):

$$P[A] = \iint_A f_{X,Y}(x',y')\,dx'\,dy' \qquad (6.1.7)$$

deci: $$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(x',y')\,dx'\,dy' \qquad (6.1.8)$$

Densitatea de probabilitate marginală a unei variabile aleatorie izolate se leagă de densitatea de probabilitate (dacă există!) a variabilei aleatorie vectoriale $\mathbf{V} = (X,Y)$ prin relaţiile:

$$f_X(x)=\int_{-\infty}^{+\infty} f_{X,Y}(x,y')\,\mathrm{d}y' \text{ şi } f_Y(y)=\int_{-\infty}^{+\infty} f_{X,Y}(x',y)\,\mathrm{d}x' \quad (6.1.9)$$

Demonstraţie:

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}=\frac{\mathrm{d}F_{X,Y}(x,+\infty)}{\mathrm{d}x}=\frac{\mathrm{d}}{\mathrm{d}x}\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty} f_{X,Y}(u,v)\mathrm{d}v\right]\mathrm{d}u$$

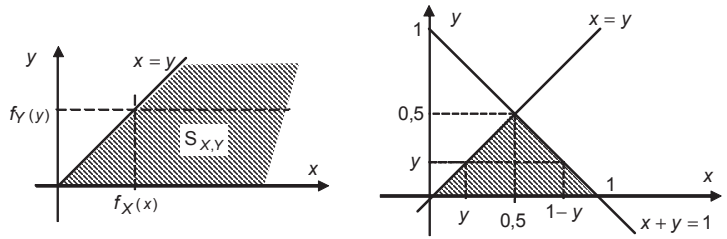
$$\underset{\frac{\mathrm{d}}{\mathrm{d}x}\int_{-\infty}^{x}\mathrm{d}u=1}{=}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,\mathrm{d}y$$

Aplicaţia 6.1.4

Se dă funcţia densitate de probabilitate de legătură:

$$f_{X,Y}(x,y)=\begin{cases} c\cdot \mathrm{e}^{-x}\cdot \mathrm{e}^{-y} & \text{pentru } 0\le y\le x<\infty \\ 0 & \text{în rest}\end{cases}$$

Să se determine: i) constanta c; ii) funcţiile densitate de probabilitate marginale; iii) $P[X+Y\le 1]$; iv) funcţia de distribuţie; v) funcţiile de distribuţie marginale.



a: Spaţiul de eşantionare corespunzător condiţiei $0\le y\le x<\infty$ **b: Regiunea corespunzătoare evenimentului** $\{X+Y\le 1\}$

Figura 6.1.4

Rezolvare:

i) Spaţiul de eşantionare al variabilei (X,Y) este ilustrat în figura 6.1.4a.

Deci: $1=\int_0^{+\infty}\int_0^{+\infty} f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x=c\int_0^{+\infty}\mathrm{e}^{-x}\int_0^{x}\mathrm{e}^{-y}\mathrm{d}y\,\mathrm{d}x=c\int_0^{+\infty}\mathrm{e}^{-x}(1-\mathrm{e}^{-x})\mathrm{d}x$, de unde rezultă că $c=2$

ii) funcţiile marginale cerute sunt:

$$f_X(x)=\int_0^{+\infty} f_{X,Y}(x,y)\mathrm{d}y=\int_0^{x}2\mathrm{e}^{-x}\mathrm{e}^{-y}\,\mathrm{d}y=2\mathrm{e}^{-x}(1-\mathrm{e}^{-x})$$

$$f_Y(y)=\int_0^{+\infty} f_{X,Y}(x,y)\mathrm{d}x=\int_y^{+\infty}2\mathrm{e}^{-x}\mathrm{e}^{-y}\,\mathrm{d}x=2\mathrm{e}^{-2y}$$

iii) Evenimentului $\{X+Y\le 1\}$ îi corespunde regiunea haşurată în imaginea 6.1.4b, iar probabilitatea solicitată este:

$$P[X+Y\le 1]=\int_0^{0,5}\int_y^{1-y}2\mathrm{e}^{-x}\mathrm{e}^{-y}\,\mathrm{d}x\,\mathrm{d}y=2\left(1-1/\sqrt{\mathrm{e}}\right)^2$$

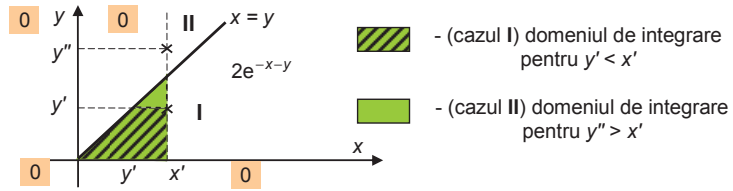


Figura 6.1.5: Expresia funcţiei $f_{X,Y}(x,y)$ pe diferite intervale

iv) Calculul funcţiei $F_{X,Y}(x,y)$ implică două cazuri distincte (figura 6.1.5):

Cazul I – pentru $0\le y'<x'<\infty$:

$$F_{X,Y}(x',y') = \int_0^{y'}\int_0^{x}2\mathrm{e}^{-x}\mathrm{e}^{-y}\mathrm{d}y\mathrm{d}x+\int_{y'}^{x'}\int_0^{y'}2\mathrm{e}^{-x}\mathrm{e}^{-y}\mathrm{d}y\mathrm{d}x$$

$$=2\int_0^{y'}\mathrm{e}^{-x}(1-\mathrm{e}^{-x})\mathrm{d}x+2\int_{y'}^{x'}\mathrm{e}^{-x}(1-\mathrm{e}^{-y})\mathrm{d}x=\left(1-\mathrm{e}^{-y'}\right)\cdot\left(1+\mathrm{e}^{-y'}-2\mathrm{e}^{-x'}\right)$$

Cazul II – pentru $0\le x'<y'<\infty$:

$$F_{X,Y}(x',y')=\int_0^{x'}\int_0^{x}2\mathrm{e}^{-x}\mathrm{e}^{-y}\mathrm{d}y\,\mathrm{d}x=\int_0^{x'}2\mathrm{e}^{-x}(1-\mathrm{e}^{-x})\,\mathrm{d}x=\left(1-\mathrm{e}^{-x'}\right)^2$$

Revenind la variabilele x şi y, rezultă că:

$$F_{X,Y}(x,y)=\begin{cases}\text{pentru } 0\le y<x<\infty: & \left(1-\mathrm{e}^{-y}\right)\cdot\left(1+\mathrm{e}^{-y}-2\mathrm{e}^{-x}\right)\\ \text{pentru } 0\le x<y<\infty: & \left(1-\mathrm{e}^{-x}\right)^2\end{cases}$$

v) $F_X(x)=F_{X,Y}(x,\infty)=\left(1-\mathrm{e}^{-x}\right)^2$ şi $F_Y(y)=F_{X,Y}(\infty,y)=1-\mathrm{e}^{-2y}$

* * *

Aplicaţia 6.1.5

Funcţia densitate de probabilitate de legătură a unei variabile aleatorie, (X,Y), gaussiană are expresia:

$$f_{X,Y}(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}}\cdot\exp\left[-\frac{x^2-2\rho\,x\,y+y^2}{2(1-\rho^2)}\right] \quad (6.1.10)$$

Să se determine: i) funcţiile densitate de probabilitate marginală;

ii) $P\left[X^2+Y^2<R^2\right]$.

Rezolvare:

i) Reprezentarea grafică a funcţiei densitate de probabilitate este prezentată în figura 6.1.6:

$$f_X(x)=\frac{\exp\left[-\frac{x^2}{2(1-\rho^2)}\right]}{2\pi\sqrt{1-\rho^2}}\cdot\int_{-\infty}^{+\infty}\exp\left[-\frac{y^2-2\rho xy}{2(1-\rho^2)}\right]\mathrm{d}y=\cdots$$

$$=\frac{\exp\left[-\frac{x^2}{2}\right]}{\sqrt{2\pi}}\cdot\int_{-\infty}^{+\infty}\frac{\exp\left[-\frac{(y-\rho x)^2}{2(1-\rho^2)}\right]}{\sqrt{2\pi(1-\rho^2)}}\mathrm{d}y=\frac{\exp\left[-\frac{x^2}{2}\right]}{\sqrt{2\pi}}$$

(integrantul este funcţia densitate de probabilitate a unei variabile aleatorie gaussiană, unidimensională de medie $\rho\cdot x$ şi varianţă $1-\rho^2$).

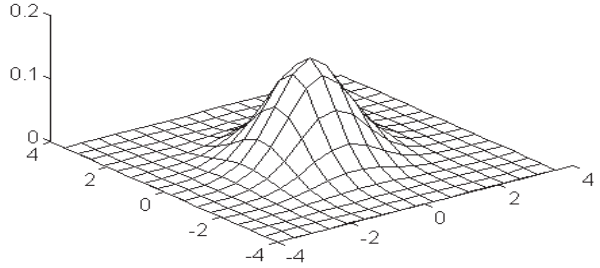


Figura 6.1.6: Funcţia densitate de probabilitate a unei variabile aleatorie gaussiană bidimensională

Datorită simetriei funcţiei $f_{X,Y}(x,y)$ avem şi: $f_Y(y)=\exp\left(-y^2/2\right)\Big/\sqrt{2\pi}$.

Observaţie: cele două funcţii densitate de probabilitate marginale urmează o lege de distribuţie gaussiană de medie 0 şi varianţă 1 .

ii) *Indicaţie*: se trece în coordonate polare.

* * *

O variabile aleatorie bidimensională mixtă (cu o componentă continuă şi una discretă), $\mathbf{V}=(X,Y)$, este descrisă de funcţia de distribuţie asociată ei:

$$F_{X,Y}(x,y)=\sum_{\{x_i:\,x_i\le x\}}P\left[X=x_i,Y\le y\right] \qquad (6.1.11)$$

6.1.1. Independenţa componentelor

Componentele X şi Y ale **unei variabile aleatorie bidimensionale** sunt *independente* dacă pentru orice evenimente unidimensionale, $A\subset \mathrm{S}_X$ şi $B\subset \mathrm{S}_Y$, se respectă relaţia:

$$P[A,B]=P[A]\cdot P[B] \qquad (6.1.1.1)$$

În general, pentru a verifica această proprietate se utilizează următoarea teoremă:

Teoremă: componentele X şi Y ale unei variabile aleatorie vectoriale bidimensionale sunt independente *dacă* şi *numai dacă*:

- ambele variabile aleatorii sunt continue:

$$f_{X,Y}(x,y)=f_X(x)\cdot f_Y(y) \qquad (6.1.1.2)$$

- cazul ambele variabile aleatorii sunt discrete:

$$p_{X,Y}(x_j,y_k)=p_X(x_j)\cdot p_Y(y_k) \qquad (6.1.1.3)$$

- cazul general:

$$F_{X,Y}(x,y)=F_X(x)\cdot F_Y(y) \qquad (6.1.1.4)$$

Aplicaţia 6.1.1.1

Verificaţi proprietăţile de independenţă ale variabilelor aleatorii din: i) aplicaţia 6.1.1; ii) aplicaţia 6.1.3; iii) aplicaţia 6.1.4; iv) aplicaţia 6.1.5.

Rezolvare:

i) funcţiile de probabilitate marginale sunt:

$$p_Q(q)=P[Q=q]=\left(1-p^M\right)\cdot\left(p^M\right)^q \text{ şi } p_R(r)=P[R=r]=\frac{1-p}{1-p^M}\cdot p^r$$

Produsul lor: $p_Q(q)\cdot p_R(r)=\cdots=(1-p)\cdot p^{M\cdot q+r}=p_{Q,R}(q,r)$, pentru orice valoare $q=0,1,\ldots$ şi $r=\overline{0,M-1}$, verifică teorema anterioară, deci variabilele aleatorii Q şi R sunt independente.

ii) $F_X(x)\cdot F_Y(y)=(1-\mathrm{e}^{-\alpha\cdot x})\cdot(1-\mathrm{e}^{-\beta\cdot x})=F_{X,Y}(x,y)$ pentru orice $x\ge 0$ şi $y\ge 0$, deci variabilele aleatorii X şi Y sunt independente.

iii) $f_X(x)=2\mathrm{e}^{-x}(1-\mathrm{e}^{-x})$, pentru $0\le x<\infty$, şi

$f_Y(y)=2\mathrm{e}^{-2y}$, pentru $0\le y<\infty$, deci:

$f_X(x)\cdot f_Y(y)=4\,\mathrm{e}^{-x}(1-\mathrm{e}^{-x})\cdot\mathrm{e}^{-y}$, pentru $0\le x<\infty$ şi $0\le y<\infty$.

Deoarece produsul $f_X(x)\cdot f_Y(y)$ este diferit de 0 şi în afara regiunii $y<x$ unde $f_{X,Y}(x,y)=0$, tragem concluzia că cele două variabile aleatorii X şi Y nu sunt independente (proprietate intuită şi datorită restricţiei impuse de relaţia $y<x$).

iv) $f_X(x)\cdot f_Y(y)=\left\{2\pi\cdot\exp\left[2(x^2+y^2)\right]\right\}^{-1}$, pentru $x,y\in\mathrm{R}$.

Pentru $\rho=0$ avem $f_{X,Y}(x,y)=f_X(x)\cdot f_Y(y)$, ceea ce înseamnă că variabilele aleatorii X şi Y sunt independente **dacă şi numai dacă** $\rho=0$. Se precizează aici că ρ este coeficientul de corelaţie între X şi Y, noţiunea aceasta fiind definită mai târziu în cuprinsul acestui capitol.

* * *

Teoremă: Dacă X şi Y sunt variabile aleatorii independente, atunci orice pereche de variabile aleatorii $g(X)$ şi $h(Y)$, definite prin funcţiile deterministe $g(*)$ şi $h(*)$, sunt independente.

Demonstraţie: Fie următoarele evenimente $A \subset \mathrm{S}_X$ şi $B \subset \mathrm{S}_Y$ pentru care avem evenimentele echivalente: $A' = \{g(x) \mid x \in A\}$ şi $B' = \{h(x) \mid x \in B\}$. Deci:

$$P[g(X) \in A', h(Y) \in B'] = P[X \in A, Y \in B] \overset{X,Y \text{ independente}}{=} P[X \in A] \cdot P[Y \in B] = P[g(x) \in A'] \cdot P[h(y) \in B']$$

relaţie ce verifică definiţia independenţei.

6.1.2. Probabilităţi condiţionate şi mărimi caracteristice condiţionate

➢ Dacă X şi Y sunt **componente discrete** ale unei variabile aleatorie bidimensionale atunci probabilitatea elementară a lui Y, condiţionată de realizarea $X = x_k$ este:

$$p_Y(y_j \mid x_k) = P[Y = y_j \mid X = x_k] = \frac{P[Y = y_j, X = x_k]}{P[X = x_k]} \quad (6.1.2.1)$$

Pentru consistenţa formulei, definim $p_Y(y_j \mid x_k)$ pentru orice x_k cu proprietatea $P[X = x_k] = 0$.

➢ Dacă X este **componenta discretă** şi Y este **componenta continuă** a unei variabile aleatorie bidimensionale atunci distribuţia lui *Y* condiţionată de realizarea $X = x_k$ este:

$$F_Y(y \mid x_k) = \frac{P[X = x_k, Y \le y]}{P[X = x_k]} \text{ pentru } P[X = x_k] > 0 \quad (6.1.2.2)$$

Dacă derivata există, atunci densitatea de probabilitate condiţionată corespunzătoare este: $f_Y(y \mid x_k) = \mathrm{d}F_Y(y, x_k)/\mathrm{d}y$

➢ Dacă X şi Y sunt **componente continue** ale unei variabile aleatorie bidimensionale, distribuţia condiţionată se defineşte printr-un procedeu de aducere la limită:

$$F_Y(y|x) = \lim_{h \to 0} P[Y \le y \mid x < X \le x + h] = \lim_{h \to 0} \frac{P[x < X \le x + h, Y \le y]}{P[x < X \le x + h]} = \lim_{h \to 0} \frac{\int_{-\infty}^{y} \int_{x}^{x+h} f_{X,Y}(u,v)\,\mathrm{d}u\,\mathrm{d}v}{\int_{x}^{x+h} f_X(u)\,\mathrm{d}u}$$

Considerând variaţia lui $f_{X,Y}(u,v)$, după u pe intervalul $(x, x+h)$, nesemnificativă, rezultă:

$$F_Y(y \mid x) = \lim_{h \to 0} \frac{\int_{-\infty}^{y} h \cdot f_{X,Y}(u,v)\,\mathrm{d}v}{f_X(x) \cdot h} = \frac{\int_{-\infty}^{y} f_{X,Y}(x,u)\,\mathrm{d}u}{f_X(x)} \quad (6.1.2.3)$$

• Dacă derivata există, **densitatea de probabilitate condiţionată**, corespunzătoare este:

$$f_Y(y \mid x) = \frac{\mathrm{d}\{F_Y(y \mid x)\}}{\mathrm{d}y} = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad (6.1.2.4)$$

Această relaţie reprezintă probabilitatea ca variabila aleatorie Y să se găsească în intervalul infinitezimal $(y, y + \mathrm{d}y)$, când X se găseşte în intervalul infinitezimal $(x, x + \mathrm{d}x)$. Ea corespunde regulei Bayes aplicată în cazul continuu.

• Dacă X este continuă şi Y discretă, probabilitatea elementară condiţionată se obţine în mod asemănător:

$$p_Y(y_k \mid x) = P[Y = y_k \mid X = x] = \lim_{h \to 0} \frac{P[Y = y_k] \,\&\, x < X \le x + h]}{P[x < X \le x + h]} = \lim_{h \to 0} \frac{\int_{x}^{x+h} f_{X,Y}(x', y_k)\,\mathrm{d}x'}{\int_{x}^{x+h} f_X(x')\,\mathrm{d}x'} = \frac{f_{X,Y}(x, y_k)}{f_X(x')} \quad (6.1.2.5)$$

Teorema probabilităţii totale (versiunea pentru variabile aleatorii vectoriale bidimensionale):

i) în cazul X discret şi Y continuu/discret:

$$P[Y \in A] = \sum_{x_k} P[Y \in A \mid X = x_k] \cdot P[X = x_k] \quad (6.1.2.6)$$

ii) în cazul X continuu şi Y continuu/discret:

$$P[Y \in A] = \int_{-\infty}^{+\infty} P[Y \in A \mid X = x] \cdot f_X(x)\,\mathrm{d}x \quad (6.1.2.7)$$

Datorită avantajelor oferite, calculul probabilităţilor unor evenimente complicate recurge deseori la această teoremă.

Aplicaţia 6.1.2.1

Numărul total de apeluri generate pe o arie geografică, X este o variabilă aleatorie Poisson de medie α. Presupunând că fiecare apel are probabilitatea p de a apărea într-o regiune specificată, R, şi că localizarea fiecărui apel este independentă de localizarea celorlalte apeluri, să se determine funcţia de probabilitate a numărului de apeluri, Y, ce apar în regiunea R.

Rezolvare: Conform teoremei probabilităţii totale se scrie că:

$$P[Y = j] = \sum_{k=0}^{+\infty} P[Y = j, X = k] = \sum_{k=0}^{+\infty} P[Y = j \mid X = k] \cdot P[X = k],$$

unde evenimentul $\{Y = j \mid X = k\}$ semnifică faptul că:

$\{$din cele k apeluri generate în reţea, j provin din regiunea $R\}$.

Conform ipotezei: $P\left[Y=j \middle| X=k\right]=\mathrm{C}_k^j\cdot p^j\cdot(1-p)^{k-j}$, deci:

$$P[Y=j] = \sum_{k=j}^{\infty}\mathrm{C}_k^j\cdot p^j\cdot(1-p)^{k-j}\cdot\frac{\alpha^k}{k!}\mathrm{e}^{-\alpha}$$

$$=\frac{(\alpha\cdot p)^j\mathrm{e}^{-\alpha}}{j!}\sum_{k=j}^{\infty}\frac{[(1-p)\alpha]^{k-j}}{(k-j)!}=\frac{(\alpha\cdot p)^j}{j!}\cdot\mathrm{e}^{-\alpha\cdot p}$$

ceea ce înseamnă că variabila aleatorie Y urmează tot o distribuţie Poisson de medie $\alpha\cdot p$.

* * *

Aplicaţia 6.1.2.2

Numărul de clienţi, ce sosesc într-un sistem de servire în intervalul de timp t, este o variabilă aleatorie ce urmează o distribuţie Poisson de parametru βt. Durata T de servire a fiecărui client este o variabilă aleatorie ce urmează o distribuţie exponenţială de parametru α.

Determinaţi funcţia de probabilitate a *numărului* N *de clienţi*, *ce sosesc pe durata* T *de servire a unui anumit client.* Se presupune că sosirile clienţilor sunt independente de timpul de servire a clientului.

Rezolvare:

Ne găsim în situaţia: o variabilă aleatorie discretă condiţionată de o variabilă aleatorie continuă. Conform teoremei probabilităţii totale:

$$P[N=k] = \int_0^{\infty}P\left[N=k\middle|T=t\right]\cdot f_T(t)\,\mathrm{d}t$$

$$=\int_0^{+\infty}(\beta t)^k/k!\cdot\mathrm{e}^{-\beta t}\cdot\alpha\cdot\mathrm{e}^{-\alpha t}\mathrm{d}t=(\beta t)^k/k!\cdot\int_0^{+\infty}t^k\cdot\mathrm{e}^{-(\alpha+\beta)\cdot t}\mathrm{d}t$$

Facem substituţia: $r=(\alpha+\beta)\cdot t$ şi deci:

$$P[N=k]=\frac{\alpha\cdot\beta^k}{k!\cdot(\alpha+\beta)^{k+1}}\underbrace{\int_0^{\infty}r^k\cdot\mathrm{e}^{-r}\mathrm{d}r}_{\Gamma(k+1)=k!}=\frac{\alpha\cdot\beta^{\,k}}{(\alpha+\beta)^{k+1}}=\frac{\alpha}{\alpha+\beta}\left(1-\frac{\alpha}{\alpha+\beta}\right)^k$$

Observaţie: variabila aleatorie N aparţine vectorului (N,T) şi urmează şi ea o lege geometrică de parametru $p=\alpha/(\alpha+\beta)$ (probabilitatea succesului într-o încercare).

* * *

Media condiţionată a variabilei aleatorie Y, când $X=x$, este:

- pentru Y continuă:

$$\mathrm{E}\left[Y\middle|X=x\right]=\int_{-\infty}^{+\infty}y\cdot f_Y(y|x)\,\mathrm{d}y \qquad (6.1.2.8)$$

- pentru Y discretă:

$$\mathrm{E}\left[Y\middle|X=x\right]=\sum\nolimits_{y_j}y_j\cdot P\left[Y=y_j\middle|X=x\right] \qquad (6.1.2.9)$$

Aplicaţia 6.1.2.3

Determinaţi media variabilelor aleatorii Y din *Aplicaţia 6.1.2.1*, utilizând formula mediei condiţionate.

Rezolvare: $\mathrm{E}[Y]=\sum_{k=0}^{+\infty}\mathrm{E}\left[Y\middle|X=k\right]\cdot P[X=k]$.

Pentru $X=x$, variabila aleatorie Y urmează o distribuţie binomială de parametrii k şi p şi deci: $\mathrm{E}[Y]=\sum_{k=0}^{+\infty}k\cdot p\cdot P[X=k]=p\cdot[X]=\alpha p$ (X fiind o variabilă aleatorie Poisson de medie α).

Observaţie: rezultatul verifică faptul că variabila aleatorie Y izolată urmează o lege de distribuţie Poisson de medie $\alpha\cdot p$.

* * *

Aplicaţia 6.1.2.4

Determinaţi media şi varianţa numărului de clienţi, N, ce sosesc pe durata de servire, T, a unui anumit client din *Aplicaţia 6.1.2.2.*

Rezolvare: Deoarece variabila aleatorie $(N\middle|T=t)$ urmează o distribuţie Poisson de parametru βt, înseamnă că:

$$\mathrm{E}\left[N\middle|T=t\right]=\beta t \text{ şi } \mathrm{E}\left[N^2\middle|T=t\right]=(\beta t)^2+\beta t, \text{ deci:}$$

- $\mathrm{E}[N]=\int_0^{+\infty}\mathrm{E}\left[N\middle|T=t\right]\cdot f_T(t)\,\mathrm{d}t=\int_0^{+\infty}\beta t\cdot f_T(t)\,\mathrm{d}t=\beta\cdot\mathrm{E}[T]$
- $\mathrm{E}\left[N^2\right]=\int_0^{+\infty}\mathrm{E}\left[N^2\middle|T=t\right]f_T(t)\,\mathrm{d}t=\int_0^{+\infty}(\beta t+\beta^2t^2)\,f_T(t)\,\mathrm{d}t=\beta\,\mathrm{E}[T]+\beta^2\,\mathrm{E}\left[T^2\right]$
- $\mathrm{VAR}[N] = \mathrm{E}\left[N^2\right]-\mathrm{E}^2[N]=\beta^2\,\mathrm{E}\left[T^2\right]-\beta^2\,\mathrm{E}^2[T]+\beta\,\mathrm{E}[T]$
  $=\beta^2\cdot\mathrm{VAR}[T]+\beta\cdot\mathrm{E}[T]$

Observaţii:

- dacă T este constant, adică $\mathrm{E}[T]=\text{constant}$ şi $\mathrm{VAR}[T]=0$, atunci variabila aleatorie N urmează o distribuţie Poisson de parametru $\beta\cdot\mathrm{E}[T]$.

- dacă T este variabilă, varianţa sa influenţează varianţa variabilei aleatorie N, amplificând-o.

- în cazul pentru care T respectă o distribuţie exponenţială de parametru α, atunci $\mathrm{E}[T]=1/\alpha$ şi $\mathrm{VAR}[T]=1/\alpha^2$, deci: $\mathrm{E}[N]=\beta/\alpha$ şi $\mathrm{VAR}[N]=\beta^2/\alpha^2+\beta/\alpha$.

Aceste rezultate sunt în concordanţă cu observaţia din *Aplicaţia 6.1.2.2.* conform căreia variabila aleatorie N urmează o lege geometrică de parametru $p=\alpha/(\alpha+\beta)$:

$$\mathrm{E}[N]=\frac{1-p}{p}=\frac{1-\alpha/(\alpha+\beta)}{\alpha/(\alpha+\beta)}=\frac{\beta}{\alpha} \text{ şi } \mathrm{VAR}[N]=\frac{1-p}{p^2}=\frac{\beta^2}{\alpha^2}+\frac{\beta}{\alpha}.$$

6.2 Variabile aleatorii *n*-dimensionale (*n* >2)

Metodele de specificare şi rezultatele obţinute pentru probabilităţile unei variabile aleatorie vectoriale bidimensionale se aplică, prin extensie, variabilelor aleatorii vectoriale *n*-dimensionale. Astfel:

- ***funcţia de distribuţie de legătură pentru n variabile aleatorii*** $X_1, X_2, \cdots, X_n$ este:

$$F_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n) = P\left[X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n\right] \quad (6.2.1)$$

- ***funcţia de distribuţie marginală*** a unora din variabilele aleatorii ce alcătuiesc un vector se obţine atribuind în cadrul funcţiilor de distribuţie (globale) valoarea ∞ argumentelor corespunzătoare celorlalte variabile aleatorii. De exemplu:

$$F_{X_1,X_2}(x_1,x_2) = F_{X_1,X_2,\cdots,X_n}(x_1,x_2,\infty,\cdots,\infty) \quad (6.2.2)$$

- ***funcţia de probabilitate*** a unui vector de *n* variabile aleatorii discrete este:

$$p_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n) = P\left[X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\right] \quad (6.2.3)$$

Cu ajutorul ei, probabilitatea oricărui eveniment, A, *n*-dimensional, se găseşte astfel:

$$P\left[(X_1,X_2,\cdots,X_n) \in A\right] = \sum_{(x_1,x_2,\cdots,x_n)\in A} p_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n)$$

- ***funcţia de probabilitate marginală a unui subansamblu***, m, de variabile aleatorii *discrete* se determină sumând după celelalte $n-m$ variabile aleatorii *discrete* din vector, pentru toate realizările posibile. De exemplu:

$$p_{X_j}(x_j) = P\left[X_j = x_j\right] = \sum_{x_1}\cdots\sum_{x_{j-1}}\sum_{x_{j+1}}\cdots\sum_{x_n} p_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n) \quad (6.2.4)$$

- ***funcţia de probabilitate a unui subansamblu*** m de variabile aleatorii *condiţionată* de restul $n-m$ variabile aleatorii ale vectorului este:

$$p_{X'_1,X'_2,\cdots,X'_n}(x'_1,\cdots,x'_m \,|\, x'_{m+1},\cdots,x'_n) = \frac{p_{X'_1,X'_2,\cdots,X'_n}(x'_1,\cdots,x'_n)}{p_{X'_{m+1},X'_2,\cdots,X'_n}(x'_{m+1},\cdots,x'_n)} \quad (6.2.5)$$

Unde $X'_1, X'_2, \cdots, X'_n$ reprezintă variabilele aleatorii $X_1, X_2, \cdots, X_n$ puse în altă ordine, iar în mod corespunzător $x'_1, x'_2, \cdots, x'_n$ reprezintă realizările $x_1, x_2, \cdots, x_n$ puse şi ele în ordinea respectivă.

Aplicaţia 6.2.1

Un sistem primeşte mesaje pe 3 linii de comunicaţie şi fie X_i numărul de mesaje recepţionate pe oră din linia i. Presupunem că funcţia de probabilitate a vectorului (X_1, X_2, X_3) este: $p_{X_1,X_2,X_3}(x_1,x_2,x_3) = \prod_{i=1}^{3}(1-a_i)a_i^{x_i}$, cu $x_i \ge 0$. Determinaţi $p_{X_1,X_2}(x_1,x_2)$ şi $p_{X_1}(x_1)$, ştiind că $0 < a_i < 1$.

Rezolvare:

$$p_{X_1,X_2}(x_1,x_2) = \sum_{x_3=0}^{+\infty} p_{X_1,X_2,X_3}(x_1,x_2,x_3)$$

$$= (1-a_1)(1-a_2)(1-a_3)a_1^{x_1}a_2^{x_2}\sum_{x_3=0}^{+\infty} a_3^{x_3} = (1-a_1)(1-a_2)a_1^{x_1}a_2^{x_2}$$

$$p_{X_1}(x_1) = (1-a_1)a_1^{x_1}$$

* * *

Variabilele aleatorii $X_1, X_2, \cdots, X_n$ sunt **simultan continue** sau, altfel spus, variabila vectorială $(X_1, X_2, \cdots, X_n)$ este continuă dacă probabilitatea oricărui eveniment *n*-dimensional, A, este dată de relaţia:

$$P\left[(X_1,\cdots,X_n)\in A\right] = \int\cdots\int_{(x_1,\cdots,x_n)\in A} f_{X_1,\cdots,X_n}(x_1,\cdots,x_n)\,dx_1\cdots dx_n \quad (6.2.6)$$

unde: $f_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n)$ este funcţia densitate de probabilitate a vectorului $(X_1, X_2, \cdots, X_n)$. Desigur că:

$$F_{X_1,\cdots,X_n}(x_1,\cdots,x_n) = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f_{X_1,X_2,\cdots,X_n}(x'_1,x'_2,\cdots,x'_n)\,dx'_1\,dx'_2\cdots dx'_n \quad (6.2.7)$$

Invers, dacă se cunoaşte funcţia densitate de distribuţie, atunci şi ea este derivabilă, adică:

$$f_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n) = \frac{\partial^n}{\partial x_1\cdots\partial x_n} F_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n) \quad (6.2.8)$$

Funcţia densitate de distribuţie marginală a unui subset de m variabile aleatorii din totalul de n este:

$$f_{X'_1,X'_2,\cdots,X'_m}(x_1,\cdots,x_m) = \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f_{X'_{m+1},X'_{m+2},\cdots,X'_n}(x_{m+1},\cdots,x_n)\,dx_{m+1}\cdots dx_n \quad (6.2.9)$$

Funcţia densitate de distribuţie a unui subset de m variabile aleatorii condiţionate de rest este (cu condiţia ca $f_{X'_{m+1},\cdots,X'_n}(x_{m+1},\cdots,x_n) > 0$):

$$f_{X'_1,\cdots,X'_n}(x_1,\cdots,x_m \,|\, x_{m+1},\cdots,x_n) = \frac{f_{X'_1,\cdots,X'_n}(x_1,\cdots,x_n)}{f_{X'_{m+1},\cdots,X'_n}(x_{m+1},\cdots,x_n)} \quad (6.2.10)$$

Aplicaţia 6.2.2

Variabila aleatorie (X_1, X_2, X_3) are funcţia densitate de probabilitate:

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3)=\frac{\exp\left[-(x_1^2+x_2^2-\sqrt{2}\cdot x_1x_2+1/2\cdot x_3^2)\right]}{2\pi\sqrt{\pi}}\text{ (Gauss)}\quad (6.2.11)$$

Determinaţi funcţia densitate de probabilitate marginală a subansamblului variabilelor aleatorii X_1 şi X_3.

Rezolvare [Ioan98]:

$$f_{X_1,X_3}(x_1,x_3)=\frac{\sqrt{\exp(-x_3^2)}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{\exp(-x_1^2-x_2^2+\sqrt{2}x_1x_2)}{2\pi/\sqrt{2}}\mathrm{d}x_1=\frac{\sqrt{\exp(-x_2^2-x_3^2)}}{2\pi}$$

* * *

Componentele $X_1,\cdots,X_n$ ale unei variabile aleatorii vectoriale sunt independente dacă:

$$P[X_1\in A_1,\cdots,X_n\in A_n]=P[X_1\in A]\cdot\ldots\cdot P[X_n\in A_n]\quad (6.2.12)$$

pentru oricare set de evenimente unidimensionale $A_1,\cdots,A_n$.

Teoremă: $X_1,\cdots,X_n$ sunt independente dacă şi numai dacă:

$$F_{X_1,\cdots,X_n}(x_1,\cdots,x_n)=F_{X_1}(x_1)\cdot\ldots\cdot F_{X_n}(x_n)\quad (6.2.13)$$

Pentru variabile aleatorii vectoriale discrete, relaţia (6.2.13) este echivalentă cu:

$$p_{X_1,\cdots,X_n}(x_1,\cdots,x_n)=p_{X_1}(x_1)\cdot\ldots\cdot p_{X_n}(x_n)\text{, pentru }\forall x_1,\cdots,x_n\ (6.2.14)$$

Pentru variabile aleatorii simultan continue, relaţia (6.2.13) este echivalentă cu:

$$f_{X_1,X_2,\cdots,X_n}(x_1,\cdots,x_n)=f_{X_1}(x_1)\cdot f_{X_2}(x_2)\cdot\ldots\cdot f_{X_n}(x_n)\quad (6.2.15)$$

Aplicaţia 6.2.3

Verificaţi independenţa variabilelor aleatorii X_1 şi X_3 din exemplul anterior (*Aplicaţia 6.2.2.*).

Rezolvare: $f_{X_1,X_3}(x_1,x_3)=\frac{\exp(-x_1^2/2)}{\sqrt{2\pi}}\cdot\frac{\exp(-x_3^2/2)}{\sqrt{2\pi}}=f_{X_1}(x_1)\cdot f_{X_3}(x_3)$

deci X_1 şi X_3 sunt variabile aleatorii independente, ce urmează o lege de distribuţie Gauss de medie zero şi varianţă unitate.

6.3 Funcţii de mai multe variabile aleatorii

Interpretarea unor fenomene aleatorii necesită deseori prelucrarea mai multor variabile aleatorii. Astfel, pe baza unui eşantion $X_1,\cdots,X_n$ alcătuit din măsurătorile repetate ale aceleiaşi mărimi supuse observaţiei se pot determina următoarele mărimi: valoarea maximă, valoarea minimă, media pe eşantion, varianţa pe eşantion, etc. Aceste mărimi, variabile aleatorii la rândul lor, se obţin aplicând diverse funcţii, $g(\cdots)$, ale eşantionului în cauză.

Funcţia de distribuţie a unei variabile aleatorie, $Z=g(X_1,X_2,\cdots,X_n)$ se află determinând echivalentul evenimentului $\{Z\le z\}$:

$$F_Z(z)=P[Z\le z]=P[\mathbf{x}\in R_Z]\text{ cu }R_Z=\{\mathbf{x}=(x_1,\cdots,x_n):g(x)=z\}\quad (6.3.1)$$

Când variabila aleatorie $(X_1,\cdots,X_n)$ este simultan continuă, atunci:

$$F_Z(z)=\int\limits_{\mathbf{x}\in R_Z}\int f_{X_1,\cdots,X_n}(x_1,x_2,\cdots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n\quad (6.3.2)$$

Relaţia (6.3.2) este valabilă şi pentru variabilele aleatorii vectoriale mixte sau discrete, caz în care o parte sau toate integralele devin sume după evenimentele din dimensiunea respectivă.

Aplicaţia 6.3.1

Fie variabila aleatorie vectorială (X,Y) şi variabila aleatorie $Z=X+Y$. Determinaţi $F_Z(z)$ şi $f_Z(z)$, considerând cunoscută funcţia densitate de probabilitate $f_{X,Y}(x,y)$.

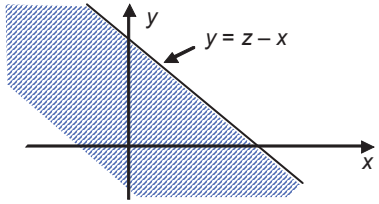


Figura 6.3.1: Regiunea $x+y\le z$

Rezolvare: Regiunea din plan, ce corespunde evenimentului $\{Z\le z\}$ este prezentată în figura 6.6.1. În acest caz:

$$F_Z(z)=\int_{-\infty}^{+\infty}\int_{-\infty}^{z-x'}f_{X,Y}(x',y')\,\mathrm{d}x'\mathrm{d}y'\text{ şi }f_Z(z)=\int_{-\infty}^{+\infty}f_{X,Y}(x',z-x')\,\mathrm{d}x'\ .$$

Dacă *X* şi *Y* sunt independente, atunci $f_{X,Y}(x,y)=f_X(x)\cdot f_Y(y)$, şi deci:

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x')\cdot f_Y(z-x')\,\mathrm{d}x'\quad (6.3.3)$$

Observaţie: relaţia (6.6.3) reprezintă **convoluţia dintre două funcţii densitate de probabilitate** (marginală, în cazul acestei aplicaţii).

6.3.1. Mărimi caracteristice ale funcţiilor de două variabile aleatorii

❖ Determinarea mărimilor aşteptate ce descriu o funcţie de mai multe variabile aleatorii urmează un procedeu similar cu cel descris în cazul unei unice variabile aleatorie. Astfel, în cazul general, **media** (valoarea aşteptată) $\mathrm{E}[Z]$ a unei funcţii Z de două variabile aleatorii X şi Y este dată de relaţii diferenţiate după tipul celor două variabile, astfel:

- pentru X şi Y continue:

$$\mathrm{E}[Z]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \qquad (6.3.1.1a)$$

- pentru X discret şi Y continuu:

$$\mathrm{E}[Z]=\sum_{x_i}\int_{-\infty}^{+\infty} g(x_i,y)\,f_{X,Y}(x_i,y)\,\mathrm{d}y \qquad (6.3.1.1b)$$

- pentru X şi Y discrete:

$$\mathrm{E}[Z]=\sum_{x_i}\sum_{y_i} g(x_i,y_i)\cdot p_{X,Y}(x_i,y_i) \qquad (6.3.1.1c)$$

Aplicaţia 6.3.1.1 **(media sumei a două variabile aleatorii)**

Fie variabila aleatorie vectorială continuă (X,Y) descrisă de funcţia densitate de probabilitate $f_{X,Y}(x,y)$.

Să se determine media variabilei aleatorie $Z=X+Y$.

Rezolvare:

$$\begin{aligned}\mathrm{E}[Z]=\mathrm{E}[X+Y] &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)\cdot f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\cdot f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\cdot f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \qquad (6.3.1.2)\\ &= \int_{-\infty}^{+\infty} x\cdot f_X(x)\,\mathrm{d}x+\int_{-\infty}^{+\infty} y\cdot f_Y(y)\,\mathrm{d}y=\mathrm{E}[X]+\mathrm{E}[Y]\end{aligned}$$

Observaţii: - Expresia (6.6.1.2) nu ţine cont dacă variabilele aleatorii individuale, X şi Y, sunt sau nu sunt independente;

- Expresia (6.6.1.2) se menţine indiferent de natura variabilei aleatorii vectoriale în cauză.

* * *

❖ **Momentul de ordinul *jk*** al variabilei aleatorie (X,Y) este:

$$\mathrm{E}\left[X^jY^k\right]=\begin{cases}\text{□ } X \text{ şi } Y \text{ continue:} & \displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^j y^k f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ \text{□ } X \text{ discret şi } Y \text{ continuu:} & \displaystyle\sum_{x_i}\int_{-\infty}^{+\infty} x_i^j y^k f_{X,Y}(x_i,y)\,\mathrm{d}y\\ \text{□ } X \text{ şi } Y \text{ discrete:} & \displaystyle\sum_{x_i}\sum_{y_n} x_i^j y_n^k p_{X,Y}(x_i,y_n)\end{cases} \qquad (6.3.1.3)$$

❖ **Corelaţia** $\mathrm{E}[X\cdot Y]$ a variabilelor aleatorii X şi Y este momentul de ordinul $j=1$ şi $k=1$. Variabilele aleatorii X şi Y sunt **ortogonale** dacă $\mathrm{E}[X\cdot Y]=0$.

❖ **Momentul centrat** de ordinul jk al variabilelor aleatorii X şi Y este momentul de ordinul jk al variabilelor aleatorii *centrate* $X-\mathrm{E}[X]$ şi $Y-\mathrm{E}[Y]$.

❖ **Covarianţa** variabilelor aleatorii X şi Y este momentul centrat de ordinul $j=1$ şi $k=1$:

$$\mathrm{COV}[X,Y]=\mathrm{E}\left[\left(X-\mathrm{E}[X]\right)\cdot\left(Y-\mathrm{E}[Y]\right)\right] \qquad (6.3.1.4)$$

Ea poate fi pusă şi sub forma:

$$\mathrm{COV}[X,Y]=\mathrm{E}[X\cdot Y]-\mathrm{E}[X]\cdot\mathrm{E}[Y] \qquad (6.3.1.5)$$

Dacă două variabile aleatorii sunt independente, covarianţa lor este nulă. Reciproca, însă, nu este adevărată decât în anumite cazuri.

❖ **Coeficientul de corelaţie**, r, al variabilelor aleatorii X şi Y este:

$$\rho=\frac{\mathrm{COV}[X,Y]}{\mathrm{STD}[X]\cdot\mathrm{STD}[Y]}=\frac{\mathrm{E}[X\cdot Y]-\mathrm{E}[X]\cdot\mathrm{E}[Y]}{\sigma_X\cdot\sigma_Y} \qquad (6.3.1.6)$$

Corelaţia, evaluată prin intermediul coeficientului de corelaţie, arată cât de "strâns" sunt legate variabilele în cauză. Astfel, este cu atât mai probabil ca două realizări particulare x şi y să fie:

- ambele peste sau sub medii, cu cât $\rho\to 1$,
- una peste şi una sub medie, cu cât $\rho\to -1$,
- lipsite de legătură, cu cât $\rho\to 0$.

Aplicaţia 6.3.1.2

Verificaţi independenţa şi calculaţi coeficientul de corelaţie dintre variabilele aleatorii X şi Y a căror distribuţie de legătură este dată în tabelul 6.3.2. Interpretaţi rezultatul.

Precizaţi o alocare de probabilităţi care să corespundă unei corelaţii puternic negative a celor două variabile.

Tabelul 6.3.2: Exemplu de distribuţie bidimensională discretă

x	y	$p_{X,Y}(x,y)$
-1	-1	7/16
-1	1	1/16
1	-1	2/16
1	1	6/16

6.3.2. Sume de variabile aleatorii

Deseori suntem confruntaţi cu probleme ce presupun contorizarea apariţiilor unui eveniment, cumularea de rezultate sau medierea unor serii de măsurători. În general, rezolvarea acestor probleme se reduce la determinarea exactă sau aproximativă a distribuţiei ce caracterizează o sumă de n variabile aleatorii.

Aplicaţia 6.3.2.1 **(media sumei a n variabile aleatorii)**

Fie variabila aleatorie vectorială continuă $(X_1, X_2, \cdots, X_n)$ descrisă de funcţia densitate de probabilitate $f_{X_1,\cdots,X_n}(x_1,\cdots,x_n)$.

Determinaţi media variabilei aleatorie $Z = X_1 + \ldots + X_n$.

Indicaţie: se aplică metoda inducţiei şi se obţine că:

$$\mathrm{E}[X_1 + X_2 + \ldots + X_n] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \ldots + \mathrm{E}[X_n] \qquad (6.3.2.1)$$

* * *

Aplicaţia 6.3.2.2 **(varianţa unei sume de două variabile aleatorii)**

Fie variabilele aleatorii X, Y şi $Z = X + Y$. Determinaţi varianţa variabilei aleatorie Z funcţie de mărimile statistice proprii variabilelor aleatorii X şi Y.

Rezolvare:

$$\mathrm{VAR}[Z] = \mathrm{E}\left[\left(Z - E[Z]\right)^2\right] = \mathrm{E}\left[\left(X + Y - \mathrm{E}[X] - \mathrm{E}[Y]\right)^2\right] =$$
$$= \mathrm{VAR}[X] + \mathrm{VAR}[Y] + 2\mathrm{COV}[X,Y]$$

* * *

Varianţa unei sume de n variabile aleatorii se obţine prin generalizarea aplicaţiei anterioare:

$$\mathrm{VAR}[X_1 + \ldots + X_n] = \sum_{k=1}^{n} \mathrm{VAR}[X_k] + \sum_{j=1 \text{ şi } j \neq k}^{n} \sum_{k=1}^{n} \mathrm{COV}[X_j, Y_k] \qquad (6.3.2.2)$$

Dacă $X_1, \cdots, X_n$ sunt variabile aleatorii independente, atunci:

$$\mathrm{VAR}[X_1 + X_2 + \cdots + X_n] = \sum_{k=1}^{n} \mathrm{VAR}[X_k] \qquad (6.3.2.3)$$

Dacă $X_1, \cdots, X_n$ sunt variabile aleatorii **independente identic distribuite** de medie μ şi varianţă σ^2 : $\mathrm{E}[X_1 + \ldots + X_n] = n\mu$ şi $\mathrm{VAR}[X_1 + \ldots + X_n] = n\sigma^2$.

Densitatea de probabilitate a unei sume S_n de variabile aleatorii continue independente se poate obţine aplicând relaţia:

$$f_{S_n}(x) = \mathrm{F}^{-1}\left\{\Phi_{X_1}(\omega) \cdot \Phi_{X_2}(\omega) \cdot \ldots \cdot \Phi_{X_n}(\omega)\right\} \qquad (6.3.2.4)$$

unde: F^{-1} este simbolul transformatei Fourier inverse, iar Φ_{X_k} este funcţia caracteristică a unei variabile aleatorii individuale X_k, pentru $k = \overline{1,n}$.

Aplicaţia 6.3.2.2

Demonstraţi relaţia (6.6.2.4).

Rezolvare: $S_n = \sum_{k=1}^{n} X_k$, şi $\Phi_{S_n}(\omega) = \mathrm{E}\left[\mathrm{e}^{j\omega S_n}\right] = \mathrm{E}\left[\mathrm{e}^{j\omega X_1} \cdot \ldots \cdot \mathrm{e}^{j\omega X_n}\right]$.

Cum $X_1, \cdots, X_n$ sunt independente, înseamnă că funcţiile de variabile aleatorii $\mathrm{e}^{j\omega X_1}, \cdots, \mathrm{e}^{j\omega X_n}$ sunt de asemenea independente, deci:

$$\Phi_{S_n}(\omega) = \mathrm{E}\left[\mathrm{e}^{j\omega X_1}\right] \cdot \ldots \cdot \mathrm{E}\left[\mathrm{e}^{j\omega X_n}\right] = \Phi_{X_1}(\omega) \cdot \ldots \cdot \Phi_{X_n}(\omega) \qquad (6.3.2.5)$$

* * *

Aplicaţia 6.3.2.3

Determinaţi funcţia densitate de probabilitate a sumei de m variabile aleatorii independente identic distribuite după o lege exponenţială de parametru α.

Rezolvare: Funcţia caracteristică a unei singure variabile aleatorii exponenţiale este: $\Phi_X(\omega) = \alpha/(\alpha - j\omega)$. Conform relaţiei (6.3.6.6) putem scrie că: $\Phi_{S_m}(\omega) = \left[\alpha/(\alpha - j\omega)\right]^m$, ceea ce reprezintă funcţia caracteristică a unei variabile aleatorii *m*-Erlang.

* * *

Aplicaţia 6.3.2.4 **(sume de variabile aleatorii în număr aleatoriu)**

Să se demonstreze că pentru o sumă, S_N, de un număr aleator N de variabile aleatorii independente identic distribuite se obţin rezultatele:

$$\mathrm{E}[S_N] = \mathrm{E}[N] \cdot \mathrm{E}[X] \text{ şi } \Phi_{S_N}(\omega) = G_N\left(\Phi_X(\omega)\right).$$

Rezolvare:

$$\mathrm{E}[S_N] = \mathrm{E}\left[\mathrm{E}\left[S_n | N\right]\right] = \mathrm{E}\left[\mathrm{E}\left[S_N | N = n\right]\right] = \mathrm{E}\left[\mathrm{E}\left[\sum_{k=1}^{n} X_k\right]\right] = \mathrm{E}\left[N \cdot \mathrm{E}[X]\right],$$

adică:

$$\mathrm{E}[S_N] = \mathrm{E}[N] \cdot \mathrm{E}[X] \qquad (6.3.2.6)$$

$$\Phi_{S_N}(\omega) = \mathrm{E}\left[\mathrm{e}^{j\omega S_N}\right] = \mathrm{E}\left[\mathrm{E}\left[\mathrm{e}^{j\omega S_N} | N\right]\right] = \mathrm{E}\left[\Phi_X(\omega)^N\right]$$
$$= \mathrm{E}\left[z^N\right]\Big|_{z = \Phi_X(\omega)} = G_N\left(\Phi_X(\omega)\right) \qquad (6.3.2.7)$$

* * *

Aplicaţia 6.3.2.5

Numărul N de cereri de serviciu adresate unui sistem de servire într-o oră este o variabilă aleatorie geometrică de parametru p, iar duratele de servire sunt variabile aleatorie independente identic distribuite după o lege exponenţială de parametru α. Determinaţi funcţia densitate de probabilitate a timpului necesar servirii cererilor primite pe durata unei ore.

Rezolvare: Funcţia generatoare a variabilei aleatorie N este $G_N(z) = p/(1 - qz)$, iar funcţia caracteristică a unei variabile aleatorie exponenţiale este $\Phi_X(\omega) = \alpha/(\alpha - j\omega)$. Conform relaţiei (6.3.6.8) funcţia caracteristică a lui S_n este: $\Phi_{S_n}(\omega) = \dfrac{p}{1 - q\alpha/(\alpha - j\omega)} = \dfrac{p(\alpha - j\omega)}{p\alpha - j\omega}$. Aplicând transformata inversă, obţinem:

$$f_{S_n}(x) = p\delta(x) + (1 - p) \cdot \exp(-p\alpha x), \text{ pentru } x \geq 0.$$

Interpretare: timpul total de servire pe durata unei ore este 0 cu probabilitatea p sau urmează o distribuţie exponenţială de medie $1/(\alpha p)$, cu probabilitatea $(1-p)$. Probabilitatea p corespunde evenimentului:

$$\{\text{pe durata unei ore nu apare nici o cerere}\},$$

iar probabilitatea $(1-p)$ corespunde evenimentului complementar.

6.3.3. Media pe eşantion

Fie X o variabilă aleatorie de medie, $\mathrm{E}[X]=\mu$, necunoscută şi n măsurători repetate, independente ale ei: $X_1, X_2, \cdots, X_n$. În aceste condiţii, funcţia de distribuţie urmată de fiecare măsurătoare este cea a variabilei aleatorie X. Cu ajutorul secvenţei $X_1, X_2, \cdots, X_n$ se construieşte **media pe eşantion** ca *estimator* al mărimii $\mathrm{E}[X]$:

$$M_n = 1/n \cdot \sum_{i=1}^{n} X_i \qquad (6.3.3.1)$$

Frecvenţa relativă, ca estimator al probabilităţii unui eveniment, reprezintă un caz particular de medie pe eşantion.

Fiind o sumă de variabile aleatorii, media pe eşantion este la rândul ei o variabilă aleatorie. Pentru a fi un bun estimator, M_n trebuie să îndeplinească următoarele condiţii:

i) în medie, să genereze valoarea corectă a parametrului estimat, $\mathrm{E}[M_n]=\mu$;

ii) să nu varieze prea mult în jurul valorii corecte ($\mathrm{E}\left[(M_n-\mu)^2\right]$ să fie *mică*).

Verificarea condiţiilor:

i) $$\mathrm{E}[M_n]=\mathrm{E}\left[1/n\cdot\sum_{i=1}^{n}X_i\right]=1/n\cdot\sum_{i=1}^{n}\mathrm{E}[X_i]=\mathrm{E}[X]=\mu \qquad (6.3.3.2)$$

(media pe eşantion este un **estimator nedeplasat** (**nedistorsionat**), a cărui medie corespunde cu parametrul sondat);

ii) $$\begin{aligned}\mathrm{E}\left[(M_n-\mu)^2\right] &= \mathrm{E}\left[(M_n-E[M_n])^2\right]=\mathrm{VAR}[M_n]\\ &=\mathrm{VAR}\left[(X_1+\cdots+X_n)/n\right]=n\cdot\mathrm{VAR}[X/n]=\sigma^2/n\end{aligned} \qquad (6.3.3.3)$$

cu $\sigma=\mathrm{VAR}[X]$.

6.3.4. Legea slabă a numerelor mari

Se presupune că: variabila aleatorie X este de medie $\mathrm{E}[X]=\mu$, necunoscută. Pentru a evalua media, se generează un *eşantion* prin repetarea experimentului şi măsurarea variabilei aleatorie X. Se obţine astfel o secvenţă $X_1, X_2, \cdots, X_n$ de variabile aleatorii i.i.d., notate generic cu X.

Pentru a estima valoarea aşteptată se calculează *media pe eşantion* :

$$M_n = 1/n \cdot \sum_{j=1}^{n} X_j \qquad (6.3.4.1)$$

Enunţul legii:

$$\lim_{n\to\infty} P\left[\,|M_n-\mu|<\varepsilon\right]=1,\ \text{cu}\ \varepsilon>0 \qquad (6.3.4.2)$$

***Demonstraţie*:**

M_n este şi ea o variabilă aleatorie, a cărei valoarea aşteptată este:

$$\mathrm{E}[M_n]=\mathrm{E}\left[1/n\cdot\sum_{j=1}^{n}X_j\right]=1/n\cdot\sum_{j=1}^{n}\mathrm{E}\left[X_j\right]=\mu\,.$$

În acest caz: $\mathrm{E}\left[(M_n-\mu)^2\right]=\mathrm{E}\left[\left(M_n-\mathrm{E}[M_n]\right)^2\right]=\mathrm{VAR}[M_n]$. Deoarece $M_n=S_n/n$, cu $S_n=X_1+\ldots+X_n$ (sumă de variabile aleatorii i.i.d.), atunci înseamnă că: $\mathrm{VAR}[M_n]=\mathrm{VAR}\left[\frac{S_n}{n}\right]=\frac{1}{n^2}\cdot\mathrm{VAR}[S_n]=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$.

Utilizând inegalitatea Cebâşev putem scrie că:

$$P\left[\,|M_n-\mathrm{E}[M_n]|\ge\varepsilon\right]\le\mathrm{VAR}[M_n]/\varepsilon^2\,,$$

şi făcând substituţiile corespunzătoare, se obţine apoi că:

$$P\left[\,|M_n-\mu|\ge\varepsilon\right]\le\sigma^2/(n\varepsilon^2)\,.$$

Luând în considerare evenimentul complementar $|M_n-\mu|<\varepsilon$, relaţia devine:

$$P\left[\,|M_n-\mu|<\varepsilon\right]=1-\sigma^2/(n\varepsilon^2) \qquad (6.3.4.3)$$

şi la limită, atunci când $n\to\infty$, avem: $\lim_{n\to\infty}P\left[\,|M_n-\mu|<\varepsilon\right]=1$, q. e. d.

Interpretare: Pentru valori suficient de mari ale lui n, media pe un eşantion a n măsurători va fi în vecinătatea valorii aşteptate, μ, cu probabilitate mare.

Aplicaţia 6.3.4.1

O tensiune constantă, dar necunoscută, trebuie estimată. Fiecare măsurătoare este perturbată de zgomot de medie 0 şi deviaţie standard $1[\mu\mathrm{V}]$. Considerând că amplitudinile zgomotului sunt variabile aleatorii independente, câte măsurători sunt necesare astfel încât, cu o probabilitate de cel puţin 0,99, media pe eşantion să fie în vecinătatea $\pm1[\mu\mathrm{V}]$ a valorii aşteptate (adevărate).

Rezolvare: Se fac notaţiile:

- amplitudinile zgomotului: N_j
- tensiunea constantă: v
- tensiunea măsurată: $X_j=v+N_j$

Se calculează: $X_j=v+N_j$: $\mathrm{E}\left[X_j\right]=v$ şi $\mathrm{VAR}\left[X_j\right]=\mathrm{VAR}\left[N_j\right]=1$.

Apoi, conform relaţiei (6.6.4.3), rezultă că:

$$1-\sigma^2/(n\varepsilon^2)\ge0{,}99\,,\ 1-1/n\ge0{,}99\ \text{şi}\ n\ge100\,.$$

6.3.5. Legea tare a numerelor mari

Se presupune că:

i) $X_1, X_2, \cdots$ este o serie de variabile aleatorii i.i.d. de medie finită $\mathrm{E}[X] = \mu$ şi varianţă finită;

ii) $M_1, M_2, \cdots$ este o serie de medii, unde M_n este media pe eşantionul ce cuprinde variabilele aleatorii $X_1, X_2, \cdots, X_{n-1}, X_n$

Enunţul legii:

$$P\left[\lim_{n\to\infty} M_n = \mu\right] = 1 \qquad (6.3.5.1)$$

Interpretare (vezi figura 6.3.5.1): cu probabilitate foarte mare, orice secvenţă particulară de medii pe eşantion se apropie de valoarea aşteptată, μ, şi rămâne în vecinătatea acesteia

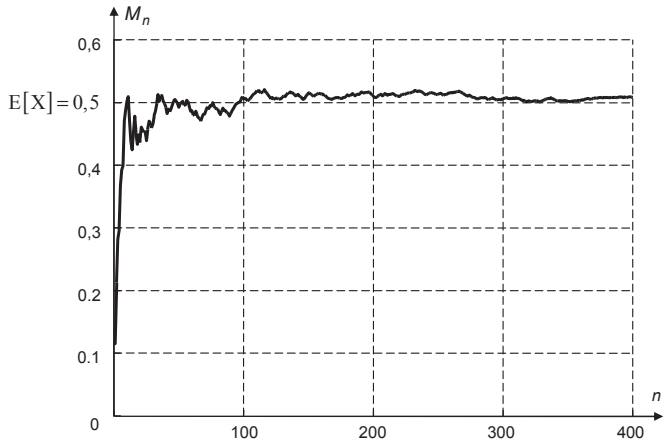


Figura 6.3.5.1: Convergenţa către $\mathrm{E}[X]$ a secvenţei de medii pe eşantion

Observaţie: Legea tare a numerelor mari demonstrează consistenţa modelelor probabilistice. Astfel, pentru fenomene ce prezintă regularitate statistică, teoria prezice convergenţa mărimilor măsurate (de exemplu mediile pe eşantion) către valoarea aşteptată ($\mathrm{E}[X] = \mu$).

Aplicaţia 6.3.5.1

Pentru a estima probabilitatea unui eveniment A, experimentul se repetă de un anumit număr de ori, n. Cât de mare trebuie să fie n astfel încât, cu probabilitatea de cel puţin 0,95, frecvenţa relativă a evenimentului să fie în vecinătatea $\pm 0{,}01$ a probabilităţii $p = P[A]$.

Rezolvare: Se ataşează evenimentului A o variabilă aleatorie X, care este funcţia indicator al succesului. Frecvenţa relativă a evenimentului este: $f_A(n) = \sum_{j=1}^{n} X_j / n$, care este o *medie pe eşantion.* Aplicăm relaţia (6.3.4.3) şi obţinem că: $P\left[\,|f_A(n) - p| < \varepsilon\right] = 1 - \sigma^2/(n\varepsilon^2)$.

Variabila aleatorie, X_j, fiind i.i.d. după o distribuţie Bernoulli, are $\mu = p$ şi $\sigma^2 = p(1-p)$. Valoarea probabilităţii p nu este cunoscută, dar, conform figurii 6.3.5.2, rezultă că: $\sigma^2 \le 1/4$. Luând în considerare valoarea maximă: $1 - \sigma^2/(n\varepsilon^2) \ge 1 - 1/(4n\varepsilon^2) \ge 0{,}95$, rezultă inegalitatea:

$1/(4n\varepsilon^2) \le 0{,}05$, adică: $n \ge 50.000$.

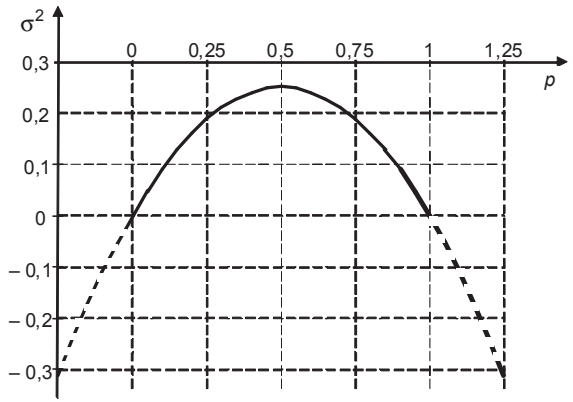


Figura 6.3.5.2: Graficul funcţiei *p* (1 – *p*)

Observaţie: Pentru acest tip de medie pe eşantion, legea tare a numerelor mari se exprimă cu relaţia: $P\left[\lim_{n\to\infty} f_A(n) = P[A]\right] = 1$.

6.3.6. Teorema limită centrală

Se presupune că:

i) $X_1, X_2, \cdots$ este o serie de variabile aleatorii i.i.d. de medie finită μ şi varianţă finită σ^2;

ii) $S_n = X_1 + X_2 + \ldots + X_n$ este suma primelor n variabile aleatorii din serie.

Enunţul teoremei:

$$\lim_{n\to\infty} P[Z_n < z] = 1/\sqrt{2\pi} \cdot \int_{-\infty}^{z} \exp\left(-x^2/2\right) \mathrm{d}x$$

unde $Z_n = \dfrac{S_n - n\mu}{\sigma\sqrt{n}}$ este variabilă aleatorie de medie 0 şi varianţă 1.

Demonstraţie: $Z_n = \dfrac{S_n - n\mu}{\sigma\sqrt{n}} = \dfrac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n}(X_k - \mu)$ şi funcţia caracteristică a variabilei aleatorie Z_n este:

$$\Phi_{Z_n}(\omega) = \mathrm{E}\left[e^{j\omega Z_n}\right] = \mathrm{E}\left[\exp\left\{\frac{j\omega}{\sigma\sqrt{n}}\sum_{k=1}^{n}(X_k-\mu)\right\}\right] = \mathrm{E}\left[\prod_{k=1}^{n}\exp\left\{\frac{j\omega}{\sigma\sqrt{n}}(X_k-\mu)\right\}\right]$$

$$= \begin{cases} \displaystyle\prod_{k=1}^{n}\mathrm{E}\left[\exp\left\{\frac{j\omega}{\sigma\sqrt{n}}(X_k-\mu)\right\}\right] & \text{- pentru variabilele aleatorii independente} \\ \left(\mathrm{E}\left[\exp\left\{\dfrac{j\omega}{\sigma\sqrt{n}}(X_k-\mu)\right\}\right]\right)^n & \text{- pentru variabilele aleatorii identic distribuite} \end{cases}$$

Dezvoltând exponenţiala, în serie, se obţine că:

$$\mathrm{E}\left[\exp\left\{\frac{j\omega}{\sigma\sqrt{n}}(X_k-\mu)\right\}\right] = \mathrm{E}\left[1+\frac{j\omega}{\sigma\sqrt{n}}(X_k-\mu)+\frac{(j\omega)^2}{2!\sigma^2 n}(X_k-\mu)^2+R(\omega)\right]$$

$$= 1+\frac{j\omega}{\sigma\sqrt{n}}\mathrm{E}[X_k-\mu]-\frac{\omega^2}{2\sigma^2 n}\mathrm{E}\left[(X_k-\mu)^2\right]+\mathrm{E}[R(\omega)]$$

$$= 1-\frac{\omega^2}{2n}+\mathrm{E}[R(\omega)]$$

Pentru valori mari ale lui n ($n \geq 20$), $\mathrm{E}[R(\omega)] << \omega^2/(2n)$, deci:

$$\Phi_{Z_n}(\omega) = \left[1-\omega^2/(2n)\right]^n \text{ şi } \lim_{n\to\to\infty}\Phi_{Z_n}(\omega) = \exp\left[-\omega^2/(2n)\right].$$

Ultima relaţie este expresia funcţiei caracteristice a unei variabile aleatorie Gauss de medie 0 şi varianţă 1, q.e.d.

Interpretare:

Funcţia de distribuţie a unei variabile aleatorie S_n ce constituie o sumă de variabile aleatorii independente, poate fi aproximată de distribuţia Gauss atunci când n devine mare.

În figurile 6.3.6.1 şi 6.3.6.2 sunt prezentate câteva exemple în acest sens.

Observaţii: Teorema este adevărată şi în următoarele cazuri:

i) variabilele aleatorii $X_1, X_2, \cdots$ sunt independente, au aceeaşi medie şi varianţă dar nu şi aceeaşi distribuţie;

ii) variabilele aleatorii $X_1, X_2, \cdots$ sunt independente şi au varianţe diferite, dar apropiate ca valoare.

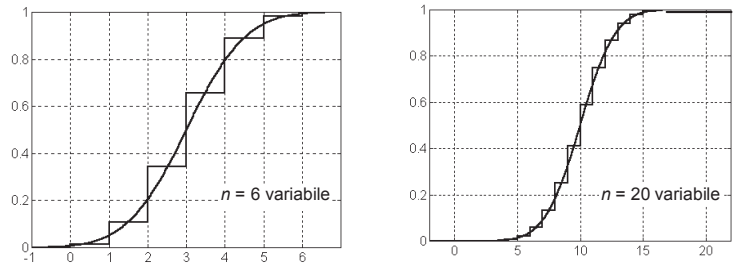


Figura 6.3.6.1: Funcţia de distribuţie a sumei de n variabile aleatorii i.i.d. după o lege Bernoulli de parametru *p* = 0.5 şi funcţia de distribuţie a unei variabile aleatorie Gauss de medie şi varianţă egale cu cele ale sumei

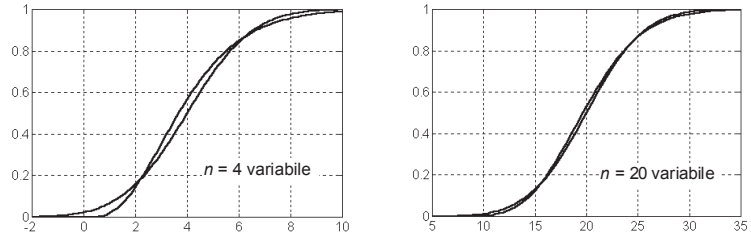


Figura 6.3.6.2: Funcţia de distribuţie a sumei de n variabile aleatorii i.i.d. după o lege exponenţială de parametru $\lambda = 1$ şi funcţia de distribuţie a unei variabile aleatorie Gauss de medie şi varianţă egale cu cele ale sumei

Aplicaţia 6.3.6.1

Se consideră timpii de servire a clienţilor variabile aleatorii i.i.d. cu media egală cu 8 secunde şi deviaţia standard egală cu 2. Să se determine:

i) probabilitatea ca durata de servire a primilor 100 de clienţi să depăşească 840 secunde;

ii) probabilitatea ca durata de servire a primilor 100 de clienţi să se găsească în intervalul $780 \div 820\,\mathrm{sec}$;

iii) care este numărul minim de clienţi serviţi, n, astfel încât, cu probabilitatea 90%, timpul servirii să depăşească 1000 secunde.

Rezolvare: i) Prin calcul exact: cele 100 de durate alcătuiesc un vector $\mathbf{X} = (X_1, X_2, \cdots, X_{100})$ de variabile aleatorii i.i.d. Suma lor este o variabilă aleatorie $S_{100} = g(\mathbf{X})$, unde funcţia *g* are expresia: $s = x_1 + x_2 + \ldots + x_{100}$.

Conform teoriei, se poate scrie că: $\mathrm{F}_{S_{100}}(s) = P[S_{100} \leq s] = P\left[\mathbf{X} \in R_{S_{100}}\right]$, unde $R_{S_{100}} = \left\{\mathbf{x} \,\middle|\, g(\mathbf{x}) \leq s\right\}$ şi $\mathbf{x} = (x_1, x_2, \cdots, x_{100})$. În cazul de faţă, ţinând cont că

variabilele aleatorii sunt independente, rezultă că:

$$F_{S_{100}}(s) = \underbrace{\int\cdots\int}_{\mathbf{x}\in R_{S_{100}}} f_{X_1}(x_1)\cdot f_{X_2}(x_2)\cdot\ldots\cdot f_{X_{100}}(x_{100})\,dx_1\,dx_2\cdots dx_{100}$$

Această expresie presupune un calcul complicat. Se justifică astfel utilizarea aproximărilor oferite de teorema limită centrală. Variabila aleatorie S_n are: media $E[S_n] = n\mu = 800$ şi varianţa $VAR[S_n] = n\sigma^2 = 400$.

Construim: $Z_{100} = (S_{100} - 800)/20$ (o variabilă aleatorie Gauss de medie 0 şi varianţă 1).

Rezultă că: $P[S_{100} > 840] \approx P[Z_{100} > (840-800)/20] = Q(2) = 0{,}0228$.

ii) $P[780 \le S_{100} \le 820] \approx P[-1 \le Z_{100} \le 1] = 1 - 2\cdot Q(2) = 0{,}682$

iii) numărul n se găseşte din relaţia: $P[S_n > 1000] = 0{,}90$.

Ştiind că $E[S_n] = 8n$ şi $VAR[S_n] = 4n$, construim variabila aleatorie $Z_n = \dfrac{S_n - 8n}{2\sqrt{n}}$. Rezultă că: $P[S_n > 1000] \approx P\left[Z_n > \dfrac{1000-8n}{2\sqrt{n}}\right] = 0{,}90$.

Conform figurii 2.6.2: $\dfrac{1000-8n}{2\sqrt{n}} = -x$, unde x se determină din tabelul 2.6.2, şi anume $x = 1{,}2815$.

Se obţine astfel ecuaţia: $8n - 1{,}2815\cdot 2\sqrt{n} - 1000 = 0$, a cărei rădăcină pozitivă este: $\sqrt{n} = 11{,}34$ sau $n = 128{,}6$.

* * *

Aplicaţia 6.3.6.2

Timpii dintre 2 sosiri succesive (de exemplu sosirea clienţilor într-un sistem cu servire) sunt variabile aleatorii i.i.d. după o lege exponenţială de medie m. Determinaţi probabilitatea ca cel de al 1000-lea eveniment să apară în intervalul $(1000 \pm 50)\cdot m$.

Rezolvare: Fie $X_1, X_2, \cdots, X_j, \cdots$ seria de timpi între două evenimente succesive şi $S_n = X_1 + X_2 + \cdots + X_n$ timpul apariţiei celui de al *n*-lea eveniment Avem: $E[X_j] = m$ şi $VAR[X_j] = m^2$, pentru $\forall j = 1,2,\cdots$;

$E[S_n] = n\cdot E[X_j] = n\cdot m$ şi $VAR[S_n] = n\cdot VAR[X_j] = n\cdot m^2$

(X_j sunt variabile aleatorii independente). Conform teoremei limită centrală:

$$P[950m \le S_{1000} \le 1050m] \approx P\left[\frac{950m - 1000m}{m\sqrt{1000}} \le Z_{1000} \le \frac{1050m - 1000m}{m\sqrt{1000}}\right]$$
$$= Q(-1{,}58) - Q(1{,}58) = 1 - 2\cdot Q(1{,}58) = 0{,}8866$$

Concluzii:

i) cu cât n creşte, cu atât mai mult variabila aleatorie S_n se găseşte în vecinătatea mediei sale cu probabilitate foarte mare;

ii) raportul n/S_n reprezintă media pe termen lung a ratei de apariţie (sosire), λ, a evenimentelor. La limită:

$$\lambda = \lim_{n\to\infty} n/S_n = 1/m \text{ evenimente/secundă.}$$

* * *

Aplicaţia 6.3.6.3

Să se rezolve *Aplicaţia 6.6.5.1* utilizând *teorema limită centrală.*

Rezolvare: Seria de variabile aleatorii i.i.d. este $I_1, I_2, \cdots, I_n, \cdots$, cu:

$$E[I_j] = p \text{ şi } VAR[I_j] = p\cdot(1-p).$$

Variabila aleatorie sumă este $f_A(n) = 1/n\cdot\sum_{j=1}^{n} I_j$, fiind caracterizată de următoarele mărimi:

- medie: $E[f_A(n)] = E[I_j] = p$;
- varianţă: $VAR[f_A(n)] = \dfrac{1}{n^2}\cdot n\cdot VAR[I_j] = \dfrac{p(1-p)}{n}$.

Variabila aleatorie Gauss corespunzătoare este: $Z_n = \dfrac{f_A(n) - E[f_A(n)]}{\sqrt{VAR[f_A(n)]}}$

şi conform teoremei limită centrală avem:

$$P[p-\varepsilon \le f_A(n) \le p+\varepsilon] \approx P\left[\frac{p-\varepsilon-p}{\sqrt{p(1-p)/n}} \le Z_n \le \frac{p+\varepsilon-p}{\sqrt{p(1-p)/n}}\right]$$
$$= 1 - 2Q\left(\frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) = 0{,}95 \text{ (valoare impusă)}$$

Din tabelul 2.6.1. se obţine că: $\dfrac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}} \approx 2$, deci $n = \dfrac{4\,p(1-p)}{\varepsilon^2}$.

Deoarece p nu este cunoscut, rezultă că:

$$n = \max_p\left[4\frac{p(1-p)}{\varepsilon^2}\right] = 4\cdot\frac{\frac{1}{4}}{\varepsilon^2} = \frac{1}{\varepsilon^2} = 10.000$$

Observaţie: Teorema limită centrală oferă o limită a lui n de 5 ori mai mică decât legea slabă a numerelor mari. Aceasta se datorează faptului că aplicarea legii slabe a numerelor mari se bazează pe o aproximaţie grosieră (Cebâşev) care, spre deosebire de teorema limită centrală, nu ţine cont de alura funcţiei densitate de probabilitate a variabilei aleatorie urmărite.

6.4 Intervale de încredere

6.4.1. Noţiuni de bază

Pentru a estima media unei variabile aleatorie, $\mathrm{E}[X]=\mu$, pe baza unui eşantion de observaţii $\mathbf{X}=(X_1, X_2, \cdots, X_n)$ se caută intervalul de valori $[u(\mathbf{X}), v(\mathbf{X})]$ în care valoarea aşteptată μ se găseşte cu o anumită probabilitate, adică:

$$P[u(\mathbf{X}) \le \mu \le v(\mathbf{X})] = 1-\alpha \qquad (6.4.1)$$

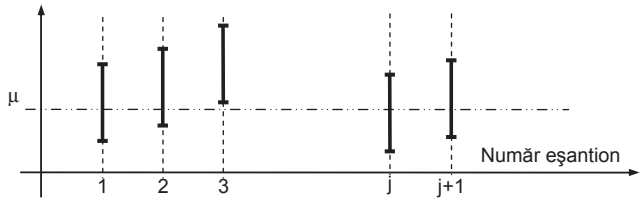


Figura 6.4.1: Secvenţă de intervale de încredere

Intervalul $[u(\mathbf{X}), v(\mathbf{X})]$ se numeşte **interval de încredere** cu probabilitatea $(1-\alpha)$, iar valoarea $1-\alpha$ se numeşte **nivel de încredere**. Conform figurii 6.4.1, nu toate intervalele de încredere obţinute conţin valoarea aşteptată μ.

Lărgimea intervalului de încredere este o măsură a acurateţei cu care este estimat parametrul de interes (μ). Cu cât intervalul este mai "strâns", cu atât parametrul este mai precis estimat.

6.4.2. Determinarea intervalelor de încredere

- **<u>Cazul 1</u>:** variabilele aleatorii $X_1, X_2, \cdots, X_n$ sunt independente şi identic distribuite după o lege Gauss de medie necunoscută şi varianţă cunoscută.

Se introduc notaţiile: $\mathrm{E}[X]=\mu$ şi $\mathrm{VAR}[X]=\sigma^2$.

Estimatorul folosit, $M_n = 1/n \cdot \sum_{j=1}^{n} X_j$, este tot o variabilă gaussiană având media $\mathrm{E}[M_n]=\mu$ şi varianţa $\mathrm{VAR}[M_n]=\sigma^2/n$. În acest caz avem:

$$P\left[-z \le \frac{M_n-\mu}{\sigma/\sqrt{n}} \le z\right] = 1-2Q(z) \text{ sau } P\left[M_n - \frac{z\sigma}{\sqrt{n}} \le \mu \le M_n + \frac{z\sigma}{\sqrt{n}}\right] = 1-2Q(z),$$

unde $\left[M_n - z\sigma/\sqrt{n}, M_n + z\sigma/\sqrt{n}\right]$ este intervalul de încredere care conţine mărimea aşteptată μ cu probabilitatea $1-2Q(z)$.

Din egalitatea $1-2Q(z)=1-\alpha$, rezultă că $Q(z_{\alpha/2})=\alpha/2$. Valorile lui $z_{\alpha/2}$, în funcţie de nivelul de încredere, sunt date în tabelul 6.4.2.1.

Pentru aceste valori, intervalul de încredere ce conţine mărimea μ cu probabilitatea $1-\alpha$ este:

$$\left[M_n - \frac{z_{\alpha/2}\cdot\sigma}{\sqrt{n}}, M_n + \frac{z_{\alpha/2}\cdot\sigma}{\sqrt{n}}\right] \qquad (6.4.2.1)$$

Tabelul 6.4.2.1: Valorile lui $z_{\alpha/2}$ la valori uzuale ale nivelului de încredere

$1-\alpha$	0,90	0,95	0,99
$z_{\alpha/2}$	1,645	1,960	2,576

<u>Aplicaţia 6.4.2.1</u>

O tensiune X are expresia $X=v+N$, în care v este o tensiune constantă, necunoscută, iar N este tensiunea de zgomot ce urmează o distribuţie Gauss de medie 0 şi varianţă $1[\mu V^2]$. Determinaţi intervalul de încredere ce conţine valoarea v cu probabilitatea 95% dacă s-a luat în considerare un eşantion de 100 de măsurători, obţinându-se media pe eşantion de $5,25[\mu V]$.

Rezolvare: Eşantionul $X_1, X_2, \cdots, X_{100}$ este alcătuit din variabile aleatorii i.i.di de medie v şi varianţă 1. Pentru $1-\alpha=0,95$, corespunde $z_{\alpha/2}=1,96$.

Intervalul de încredere cerut se determină prin particularizarea expresiei (6.4.2.1), adică: $\left(5,25-\frac{1,96\cdot 1}{10}; 5,25+\frac{1,96\cdot 1}{10}\right)=(5,05; 5,45)$.

* * *

- **<u>Cazul 2</u>:** variabilele aleatorii $X_1, X_2, \cdots, X_{100}$ sunt independente şi identic distribuite după o lege Gauss de medie şi varianţă necunoscute.

În această situaţie se determină şi varianţa eşantionului de realizări:

$$V_n^2 = \frac{1}{n-1}\cdot\sum_{j=1}^{n}\left(X_j - M_n\right)^2 \qquad (6.4.2.2)$$

Înlocuind varianţa aşteptată, σ^2, cu estimatorul ei, intervalul de încredere este:

$$\left(M_n - \frac{zV_n}{\sqrt{n}}, M_n + \frac{zV_n}{\sqrt{n}}\right) \qquad (6.4.2.3)$$

iar probabilitatea ca valoarea aşteptată, μ, să se găsească în acest interval este:

$$P\left[M_n - \frac{zV_n}{\sqrt{n}} \le \mu \le M_n + \frac{zV_n}{\sqrt{n}}\right] = P\left[-z \le \frac{M_n-\mu}{V_n/\sqrt{n}} \le z\right]$$

Termenul al doilea al egalităţii conţine variabila $W = \dfrac{M_n - \mu}{V_n/\sqrt{n}}$ ce poate fi pusă sub următoarea formă:

$$W = \frac{\sqrt{n}\cdot(M_n-\mu)/\sigma}{V_n/\sigma} = \frac{(M_n-\mu)\cdot\sqrt{n}/\sigma}{\sqrt{(n-1)^2\dfrac{V_n^2}{\sigma^2}\Big/(n-1)}}$$

Deoarece:

- numărătorul este o variabilă aleatorie Gauss de medie 0 şi varianţă 1 (M_n este variabilă aleatorie Gauss de medie μ şi varianţă σ^2/n),
- $(n-1)\dfrac{V_n^2}{\sigma^2}$ este o variabilă aleatorie χ^2 cu $(n-1)$ grade de libertate,
- numărătorul şi numitorul sunt variabile aleatorii independente (deoarece pentru variabile aleatorii independente, identic distribuite după o lege Gauss, media pe eşantion şi varianţa pe eşantion sunt variabile aleatorii independente),

variabila aleatorie W urmează distribuţia student, t, cu $(n-1)$ grade de libertate:

$$f_{n-1}(y) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\cdot\sqrt{\pi(n-1)}}\cdot\left(1+\frac{y^2}{n-1}\right)^{-n/2}.$$

Tabelul 6.4.2.2: Valorile lui $z_{\alpha/2,\,n-1}$ pentru intervalele de încredere oferite de relaţia (6.4.2.4)

$n-1$	$1-\alpha$		
	0,90	**0,95**	**0,99**
1	6,314	12,706	63,657
2	2,920	4,303	9,925
3	2,353	3,182	5,841
4	2,132	2,776	4,604
5	2,015	2,571	4,032
6	1,943	2,447	3,707
7	1,895	2,365	3,499
8	1,860	2,306	3,355
9	1,833	2,262	3,250
10	1,812	2,228	3,169
15	1,753	2,131	2,947
20	1,725	2,086	2,845
30	1,697	2,042	2,750
40	1,684	2,021	2,704
60	1,671	2,000	2,660
∞	1,645	1,960	2,576

Ţinând cont că funcţia densităţii de probabilitate $f_{n-1}(y)$ este simetrică faţă de $y=0$ se obţine:

$$P\left[M_n - \frac{z\cdot V_n}{\sqrt{n}} \le \mu \le M_n + \frac{z\cdot V_n}{\sqrt{n}}\right] = \int_{-z}^{z} f_{n-1}(y)\,\mathrm{d}y = 1 - 2F_{n-1}(-z) \quad (6.4.2.4)$$

Pentru a obţine un interval de încredere la un anumit nivel de încredere, $1-\alpha$, trebuie determinat $z_{\alpha/2,\,n-1}$ pentru care: $2F_{n-1}\left(-z_{\alpha/2,\,n-1}\right) = \alpha$.

Valorile uzuale ale lui $z_{\alpha/2,\,n-1}$ pentru calculul intervalelor de încredere sunt prezentate în tabelul 6.4.2.2.

Observaţii:

- Faptul că varianţa nu este cunoscută se reflectă şi în dispersia mai mare a distribuţiilor t faţă de cea gaussiană. Acest lucru conduce la obţinerea de intervale de încredere mai largi.
- Distribuţia Gauss aproximează distribuţia student, t, cu atât mai bine cu cât numărul de măsurători al eşantionului n creşte. Se confirmă astfel că şi în cazul variabilelor aleatorii $X_1, X_2, \cdots, X_n$ de medie şi varianţă necunoscută, pentru n mari, intervalul de încredere se determină din expresia (6.4.2.1) în care σ se înlocuieşte cu V_n. În figura 6.4.2 se compară câteva distribuţii t cu distribuţia Gauss de medie 0 şi varianţă 1.

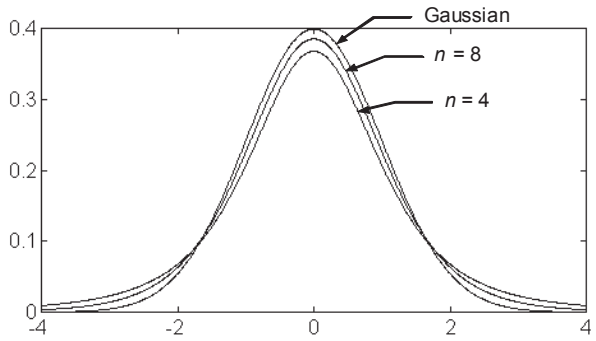


Figura 6.4.2: Funcţia densitate de probabilitate pentru distribuţia Gauss şi două distribuţii t ($n = 4$ şi $n = 8$)

Aplicaţia 6.4.2.2

Timpul de viaţă al unei tip de componente se presupune a urma o distribuţie Gauss. Opt componente sunt testate şi s-au obţinut:

- media pe eşantion = 10 zile, şi
- deviaţia standard = 2 zile.

Determinaţi intervalul de încredere pentru un nivel de încredere de 99%.

Rezolvare: Pentru $1-\alpha=0,99$ şi $n-1=7$, din tabelul 6.4.2.2 se obţine $z_{\alpha/2;\,7}=3,499$. În consecinţă, intervalul de încredere este:

$$\left(10-\frac{3,499\cdot 2}{\sqrt{8}};10+\frac{3,499\cdot 2}{\sqrt{8}}\right)=(7,53;12,47)=10\pm 2,47$$

* * *

- **Cazul 3:** variabilele aleatorii $X_1,X_2,\cdots,X_n$ urmează o distribuţie non-gaussiană de medie şi varianţă necunoscută.

Pentru a utiliza rezultatele anterioare şi în acest caz, metoda de determinare a intervalelor de încredere urmează următorul algoritm (**metoda mediilor de grup**):

- Se efectuează un număr de k experimente, fiecare cu eşantionul său de observaţii, $\mathbf{X}_k$;
- Se creează un eşantion de medii pe eşantioane : $\mathbf{M}=(M_1,M_2,\cdots,M_k)$;
- Considerând că numărul de observaţii n dintr-un eşantion X_i este suficient de mare, media pe eşantion, M_j, este o variabilă aleatorie Gauss (conform teoremei limită centrală). În acest caz, intervalul de încredere se determină conform **cazului 2**, setul de variabile aleatorii $X_1,X_2,\cdots,X_n$ fiind înlocuit cu setul de medii pe eşantion $M_1,M_2,\cdots,M_k$.

Aplicaţia 6.4.2.3

Un program de simulare generează o variabilă aleatorie distribuită după o lege exponenţială de medie necunoscută. 200 de rezultate au fost obţinute şi repartizate în 10 eşantioane (grupe) a câte 20 de rezultate fiecare. Mediile pe fiecare eşantion sunt:

1,04190 0,64064 0,80967 0,75852 1,12439
1,30220 0,98478 0,64574 1,39064 1,26890.

Determinaţi intervalul de încredere ce conţine media variabilei aleatorie cu probabilitatea 90%.

Rezolvare: Media şi varianţa pe eşantionul de medii pe grupe sunt:

$$M_{10}=0,99674 \quad \text{şi} \quad V_{10}{}^2=0,07586.$$

Pentru $1-\alpha=0,90$, corespunde $z_{\alpha/2;\,9}=1,833$ (valoarea se ia din tabelul 6.4.2.2 deoarece $n=10$ nu este suficient de mare pentru a lua în considerare aproximaţia Gauss). Se obţine intervalul de încredere $(0,83706;1,15639)$ care sugerează că $\mathrm{E}[X]\cong 1$.

6.5 Metoda fazelor

Principiul constă în a exprima în mod exact sau aproximativ o funcţie oarecare de densitate de probabilitate printr-o combinaţie, mai mult sau mai puţin complexă, de funcţii exponenţial negative [Doy89], [Les92] . Combinaţia acestora

poate fi descrisă prin intermediul unei diagrame de **stări** (numite **faze**) şi tranziţii instantanee, timpii dintre două tranziţii succesive urmând distribuţii exponenţial negativ. Descrierea exactă a unei variabile aleatorie folosind această metodă presupune că transformata sa, Laplace, este raţională, în caz contrar fiind vorba de o aproximare a cărei precizie este controlabilă.

În cele ce urmează, se prezintă patru structuri tipice, şi anume: serie, paralel, serie-paralel şi Cox. Toate acestea sunt cazuri particulare ale aşa numitelor distribuţii matriceal exponenţiale, descrise în capitolul Trafic, după familiarizarea cu instrumentarul analitic, de studiu al lanţurilor Markov.

6.5.1. Structură tip "serie"

Fie o structură tip "serie", cu reprezentarea grafică din figura 6.5.1. Cazul avut în vedere consideră că fazele de servire sunt independente şi identic distribuite şi că urmează o aceeaşi lege, de forma:

$$f_{\tau_i}(x)=k\mu\cdot\exp(-k\mu\cdot x),\ \text{cu}\ i=\overline{1,k} \qquad (6.5.1)$$

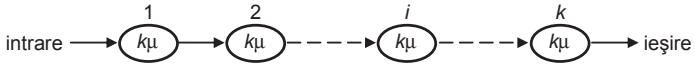


Figura 6.5.1: Structură de faze în serie

Timpul total de parcurgere a reţelei de faze, τ, în fapt o variabilă aleatorie pe care urmărim să o sintetizăm cu ajutorul acestei structuri, este suma celor k variabile i.i.d.; prin urmare, generalizând rezultatul obţinut în urma rezolvării aplicaţiei 6.6.1, se poate scrie că:

$$f_\tau(x)=f_{\tau_1}(x)\otimes f_{\tau_2}(x)\otimes\cdots\otimes f_{\tau_k}(x)$$

transformata Laplace corespunzătoare având expresia:

$$\hat{\tau}(s)=\left(\frac{k\mu}{s+k\mu}\right)^k \qquad (6.5.2)$$

cu funcţia originară:

$$f_\tau(x)=\frac{k\mu\cdot(k\mu\cdot x)^{k-1}\cdot\exp(-k\mu\cdot x)}{(k-1)!} \qquad (6.5.3)$$

Folosind formula generatoare a momentelor (2.5.14) se obţine că: $\mathrm{E}[\tau]=1/\mu$ şi $\mathrm{E}\left[\tau^2\right]=1/(k\mu^2)+1/\mu^2$, deci:

$$\mathrm{VAR}[\tau]=1/(k\mu^2) \qquad (6.5.4)$$

Prin urmare, având la dispoziţie media şi varianţa variabilei aleatorie τ, urmărite, parametrii structurii echivalente sunt daţi de relaţiile:

$$\mu=1/\mathrm{E}[\tau]\ \text{şi}\ k=\mathrm{E}^2[\tau]/\mathrm{VAR}[\tau],\ \text{cu}\ k\ \text{întreg şi}\ k\geq 1 \qquad (6.5.5)$$

Dacă se ia în considerare factorul de formă, pentru structura serie este valabilă următoarea relaţie:

$$1\leq\varepsilon\leq 2 \qquad (6.5.6)$$

Prin urmare, structura serie poate fi folosită pentru aproximarea unei variabile aleatorii **mai puţin dispersivă** decât structura exponenţială de bază, **distribuţii**le obţinute astfel numindu-se, în consecinţă, **hipoexponenţiale**. La limită, $\varepsilon = 1$, variabila aleatorie corespunde unei constante.

6.5.2. Structură tip "paralel"

Pentru o structură ca cea din figura 6.5.2, coeficienţii α_i, cu $\sum_{i=1}^{r}\alpha_i = 1$, reprezintă probabilitatea de alegere a fazei i.

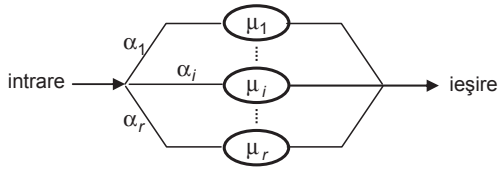


Figura 6.5.2: Structură de faze în paralel

Aplicând teorema probabilităţii totale, este clar că:

$$f_\tau(x) = \sum_{i=1}^{r}\alpha_i \mu_i \exp(-\mu_i \cdot x) \qquad (6.5.7)$$

iar transformata Laplace corespunzătoare este:

$$\hat{\tau}(s) = \sum_{i=1}^{r}\alpha_i \mu_i / (s+\mu_i) \qquad (6.5.8)$$

Se obţin, în consecinţă, pentru momentele variabilei aleatorii τ, expresiile:

$$\mathrm{E}[\tau] = \sum_{i=1}^{r}\alpha_i/\mu_i \text{ şi } \mathrm{E}\left[\tau^2\right] = 2\sum_{i=1}^{r}\alpha_i/\mu_i^2 \qquad (6.5.9)$$

de unde:

$$\varepsilon = \frac{2\sum_{i=1}^{r}\alpha_i/\mu_i^2}{\left(\sum_{i=1}^{r}\alpha_i/\mu_i\right)^2} \qquad (6.5.10)$$

Pentru două variabile aleatorii X şi Y, este valabilă inegalitatea Cauchy-Swartz, adică: $\left(\overline{X\cdot Y}\right)^2 \le \overline{X^2}\cdot\overline{Y^2}$, ceea ce, pentru cazul de faţă, se traduce în:

$\left(\sum_{i=1}^{r}\alpha_i/\mu_i\right)^2 \le \left(\sum_{i=1}^{r}\alpha_i\right)\cdot\left(\sum_{i=1}^{r}\alpha_i/\mu_i^2\right)$ şi, cum $\sum_{i=1}^{r}\alpha_i = 1$, se obţine că:

$$\left(\sum_{i=1}^{r}\alpha_i/\mu_i\right)^2 \le \sum_{i=1}^{r}\alpha_i/\mu_i^2 \qquad (6.5.11)$$

Aceasta înseamnă că de fapt:

$$2 \le \varepsilon \le \infty \qquad (6.5.12)$$

adică o **structură de tip paralel este mai dispersivă** decât structura exponenţială de bază.

Distribuţiile de tipul (6.5.7) se numesc **hiperexponenţiale de ordinul *r***, iar alegerea parametrilor implicaţi (μ_i, α_i şi r) face posibilă adaptarea lor la o gamă largă de aplicaţii experimentale. Cunoaşterea valorilor mediei şi varianţei permite scrierea următoarelor expresii de corespondenţă:

$$\mathrm{E}[\tau] \equiv \sum_{i=1}^{r}\alpha_i/\mu_i \text{ şi } \mathrm{VAR}[\tau] \equiv 2\sum_{i=1}^{r}\alpha_i/\mu_i^2 - \mathrm{E}^2[\tau] \qquad (6.5.13)$$

Aplicaţia 6.5.1

Determinaţi structura care corespunde unei variabile aleatorie de medie $40\,\mathrm{sec}$ şi deviaţie standard de i) $20\,\mathrm{sec}$, ii) $20\sqrt{10}\,\mathrm{sec}$.

6.5.3. Structură tip "serie-paralel"

Combinarea structurilor serie şi paralel (ca în figura 6.5.3) permite obţinerea unor aproximări foarte fine, dar evident cu preţul unor calcule mai laborioase.

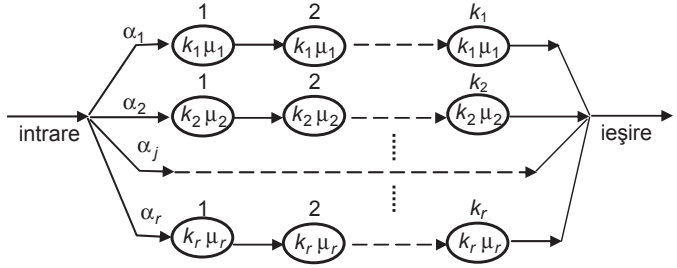


Figura 6.5.3: Structură de faze "paralel-serie"

Funcţiile de distribuţie vor fi asociate unor formule de genul (6.5.3) şi (6.5.7), având forma:

$$f_\tau(x) = \sum_{i=1}^{r}\alpha_i \frac{k_i\mu_i\cdot(k_i\mu_i\cdot x)^{k_i-1}\cdot\exp(-k_i\mu_i\cdot x)}{(k_i-1)!} \qquad (6.5.14)$$

şi respectiv transformata:

$$\hat{\tau}(s) = \sum_{i=1}^{r}\alpha_i\left(\frac{k_i\mu_i}{s+k_i\mu_i}\right)^{k_i} \qquad (6.5.15)$$

Procedând ca şi pentru variantele anterioare, se poate demonstra că structurile serie-paralel au un factor de formă cu valori supraunitare, $1 \le \varepsilon \le \infty$, acoperind deci toate cazurile posibile.

Gradul de generalizare al structurii din figura 6.5.3 se poate extinde dacă pe fiecare ramură j se diferenţiază fazele prin valorile medii asociate, $1/\mu_{j,i}$, ceea ce va influenţa corespunzător expresia funcţiei de distribuţie:

$$\hat{\tau}(s)=\sum_{j=1}^{r}\alpha_i\prod_{i=1}^{k}\frac{\mu_{j,i}}{s+\mu_{j,i}}=\sum_{j=1}^{r}\alpha_i\prod_{i=1}^{k}\hat{\tau}_{j,i}(s) \qquad (6.5.16)$$

cu $\hat{\tau}_{i,j}(s)=$transformata funcţiei de distribuţie corespunzătoare fazei i de pe ramura j a structurii în discuţie.

6.5.4. Structură tip Cox

Se foloseşte o structură serie-paralel de forma celei din figura 6.5.4, cu respectarea evident a condiţiei $\alpha_i+\beta_i=1$.

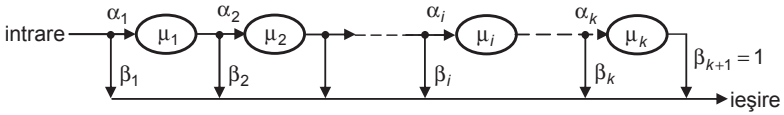


Figura 6.5.4: Structură de tip "Cox"

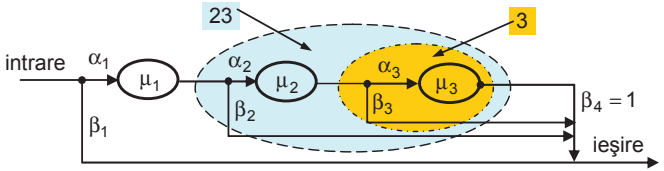


Figura 6.5.5: Structură "Cox" cu 3 faze

Stabilirea expresiei transformatei $\hat{\tau}(s)$ se face iniţial doar pentru o structură restrânsă la 3 celule (figura 6.5.5), cu următoarele precizări:

- subreţeaua 3 este o structură paralelă de ordinul 2, pentru care cea de a doua fază are media de valoare 0 ($1/\mu=0$). În consecinţă:

$$\hat{\tau}_3(s)=\beta_3+\alpha_3\mu_3/(s+\mu_3) \qquad (6.5.17)$$

- subreţeaua 23 este o structură serie-paralel cu 2 braţe, dintre care unul corespunde unei faze cu μ_2, în serie cu o subreţea 3, iar celălalt braţ corespunde unei faze de medie 0:

$$\hat{\tau}_{23}(s)=\beta_2+[\alpha_2\mu_2/(s+\mu_2)]\cdot\hat{\tau}_3(s) \qquad (6.5.18)$$

Deci pentru structura globală expresia este:

$$\hat{\tau}(s)=\beta_1+[\alpha_1\mu_1/(s+\mu_1)]\cdot\{\beta_2+\alpha_2\mu_2/(s+\mu_2)\cdot[\beta_3+\alpha_3\mu_3/(s+\mu_3)]\} \qquad (6.5.19)$$

care poate fi acum generalizată pentru *k* celule, căpătând forma:

$$\hat{\tau}(s)=\beta_1+\sum_{i=1}^{k}\beta_{i+1}\prod_{j=1}^{i}\alpha_j\mu_j/(s+\mu_j),\ \text{cu}\ \beta_{k+1}=1 \qquad (6.5.20)$$

Cu ajutorul formulei generatoare a momentelor se pot evalua primele două momente ale variabilei aleatorii căutate, şi anume:

$$E[\tau]=\sum_{i=1}^{k}\alpha_1\alpha_2\cdots\alpha_i/\mu_i \ \text{şi}\ VAR[\tau]=\sum_{i=1}^{k}\alpha_1\alpha_2\cdots\alpha_i/\mu_i^{\,2} \qquad (6.5.21)$$

În ceea ce priveşte factorul de formă, structura Cox acoperă spaţiul $\varepsilon\geq 1$.

Aplicaţia 6.5.2

Considerând că procesarea cererilor într-un sistem de servire presupune desfăşurarea mai multor activităţi, în conformitate cu o structură Cox cu:

$$\alpha_1=1,\ \alpha_2=0{,}4,\ \alpha_3=0{,}5,\ \beta_4=1 \ \text{şi}\ 4\mu_1=2\mu_2=\mu_3=8\text{msec}^{-1},$$

stabiliţi cu ce şansă timpul de servire a unei cereri depăşeşte $0{,}4\text{msec}$.

Indicaţie: Transformata Laplace se scrie sub formă de fracţii simple.

capitolul 7

Generarea numerelor aleatorii

Pentru a genera numere aleatorii se pot utiliza diverse procedee. Dintre acestea, în practică, s-au impus:

- folosirea tabelelor de numere aleatorii întocmite pe baza unor observaţii îndelungate ale unui fenomen aleatoriu,
- utilizarea directă a unui fenomen fizic aleatoriu,
- aplicarea metodelor matematice.

În simularea cu ajutorul calculatorului, cele mai utilizate sunt procedeele aritmetice implementate cu coduri, numite **generatoare de numere aleatorii** (sau secvenţe) **GNA**, [Ioan]. Tehnicile de simulare ce folosesc GNA se mai numesc "Tehnici Monte Carlo". Pentru a acoperi o plajă largă de aplicaţii, un program de simulare trebuie să fie înzestrat cu un ansamblu de astfel de generatoare "specializate" pe diferite repartiţii. Dintre acestea, cel ce generează numere distribuite uniform în intervalul $[0,1]$ ocupă o poziţie privilegiată, deoarece este implicat în funcţionarea tuturor celorlalte generatoare.

7.1 Metode de generare a numerelor aleatorii distribuite uniform

Acurateţea, cu care un calculator poate modela comportamentul unei variabile aleatorii *continue*, este limitată de:

- precizia şi domeniul de reprezentare a numerelor, care permite doar mulţimi finite pentru spaţiul realizărilor, S_X (de exemplu pentru un calculator ce reprezintă un întreg fără semn pe m biţi, $S_X = \{1,2,\cdots,M\}$, unde $M = 2^m - 1$; împărţind aceste numere la M se obţine un nou spaţiu de eşantionare care "acoperă", mai mult sau mai puţin, intervalul $[0,1]$).
- eficienţa calculului per realizare: este asigurată doar de utilizarea formulelor **recursive**. Acestea produc, însă, ***secvenţe pseudoaleatorii*** (periodice).

Ca exemplu de metodă **nerecursivă** se poate considera repetarea de m ori a unui experiment de tip Bernoulli, fiecare încercare corespunzând unui bit dintr-un număr binar de m biţi. Cu toate aspectele pozitive privind proprietăţile statistice ce le oferă, astfel de procedee nu sunt utilizate în cadrul programelor de simulare, datorită timpului excesiv consumat de operaţiile implicate.

Pentru generarea de numere distribuite uniform între 1 şi M se utilizează frecvent **metoda puterii reziduului** (congruenţială) ce utilizează relaţia:

$$X(k) = [\alpha \cdot X(k-1)] \bmod M\ ,\ k = 1,2,3,\cdots \qquad (7.1.1)$$

unde: - α este un număr între 1 şi M de a cărui alegere depinde lungimea secvenţei generate;
- M este un număr prim, mare sau o putere a unui număr prim.

Acest generator de numere aleatorii are ca ***origine*** (***seed***) realizarea $X(0)$ care este oferită de utilizator ($X(0) = 0$ nu este permisă deoarece acest lucru conduce la generarea continuă de zerouri!). Secvenţa generată fiind periodică, valoarea originii (numită şi **punct de plecare**) precizează din ce poziţie a acesteia începe să funcţioneze generatorul.

Aplicaţia 7.1.1

Determinaţi secvenţa generată de formula (7.1.1) dacă:

i) $M = 11$, $\alpha = 7$, $X(0) = 1$;

ii) $M = 11$, $\alpha = 3$, $X(0) = 3$;

iii) $M = 4$, $\alpha = 2$, $X(0) = 1$;

Rezolvare:

i) $X(1) = (7 \cdot 1) \bmod 11 = 7$; $X(2) = (7 \cdot 1) \bmod 11 = 5$ ş.a.m.d..

Secvenţa generată este: $\underset{\downarrow\ X(0)}{\cdots,1,7,5,2,3,10,4,6,9,8,1,7,5,2,3,1,0,4,6,9,8,1,7,5,\cdots}$

De remarcat că periodicitatea secvenţei este maximă, cuprinzând toate valorile între 1 şi 10.

ii) $X(1) = (3 \cdot 3) \bmod 11 = 9$; $X(2) = (3 \cdot 9) \bmod 11 = 5$ ş.a.m.d..

Secvenţa generată este: $\underset{\downarrow\ X(0)}{\cdots 1,3,9,5,4,1,3,9,5,4,1,3\cdots}$

Din punct de vedere statistic, acest generator este nesatisfăcător

deoarece perioada sa omite o bună parte din valorile admise ($2, 6, 7, 8, 10!$).

iii) $X(1)=(2\cdot 1)\bmod 4=2$; $X(2)=(2\cdot 2)\bmod 4=4\,0$; $X(3)=(2\cdot 0)\bmod 4=0$

etc. Secvenţa generată este: $1, 2, 0, 0, 0,...$

* * *

Analiza proprietăţilor statistice ale formulei recursive (7.1.1), a arătat că, pentru diferite valori ale parametrilor α şi M, există un număr restrâns de cazuri care generează secvenţe satisfăcătoare. Astfel, pentru a obţine secvenţe de perioadă maximă este necesar ca α să fie *rădăcină primitivă* a lui M.

Analiza proprietăţilor statistice ale secvenţelor generate de formula (7.1.1), pentru diferite valori ale parametrilor α şi M au reţinut un număr restrâns de cazuri satisfăcătoare. Pentru calculatoarele care memorează numerele întregi pe 32 de biţi (31 pentru valoarea absolută şi 1 pentru semn) studii extensive au identificat următoarea soluţie:

$$X(k)=\left[16807\cdot X(k-1)\right]\bmod(2^{31}-1) \qquad (7.1.2)$$

Această formulă generează o secvenţă de numere independente distribuite uniform între 1 şi $2^{31}-1$, de perioadă maximă $2^{31}-1\cong 10^9$.

Pentru a genera valori în intervalul $[0,1]$ se utilizează expresia:

$$U(k)=\frac{X(k)}{M} \qquad (7.1.3)$$

Compilatoarele utilizate pe calculatoare sunt înzestrate cu astfel de GNA optimizate în funcţie de numărul de biţi (lungimea cuvântului) utilizaţi pentru a memora un întreg şi de modul de realizare a operaţiilor aritmetice. Pe lângă aceste GNA-uri, dependente de maşină, există şi soluţii independente de maşină ("portabile") care oferă perioade lungi şi proprietăţi statistice foarte bune. Din această categorie face parte şi algoritmul **Witchmann-Hill**. Principial, el utilizează simultan 3 formule de recurenţă de tipul expresiei (7.1.2):

$$\begin{aligned} X(k)&=\left[171\cdot X(k-1)\right]\bmod 30269\\ Y(k)&=\left[172\cdot Y(k-1)\right]\bmod 30307\\ Z(k)&=\left[170\cdot Z(k-1)\right]\bmod 30323 \end{aligned} \qquad (7.1.4)$$

Aplicarea directă a acestor formule cere ca tipul de variabilă *întreg* (`int`), oferit de compilator să cuprindă valoarea 5212632. Numărul generat la pasul k este obţinut prin mediere, conform expresiei următoare:

$$U(k)=\frac{X(k)/30269+Y(k)/30307+Z(k)/30323}{3} \qquad (7.1.5)$$

Se recomandă ca, în această expresie, operaţiile să se efectueze în virgulă mobilă, dublă precizie.

În cazul în care compilatorul permite înmagazinarea de numere întregi până la 32767, relaţiile (7.1.4) sunt înlocuite cu:

$$\begin{aligned} X(k)&=171\cdot\left[X(k-1)\right]\bmod 177-2\cdot\left[X(k-1)/177\right]\\ Y(k)&=172\cdot\left[Y(k-1)\right]\bmod 176-2\cdot\left[Y(k-1)/176\right]\\ Z(k)&=170\cdot\left[Z(k-1)\right]\bmod 178-2\cdot\left[Z(k-1)/178\right]\\ &\text{Dacă } X(k)<0:\quad X(k)=X(k)+30269\\ &\text{Dacă } Y(k)<0:\quad Y(k)=Y(k)+30307\\ &\text{Dacă } Z(k)<0:\quad Z(k)=Z(k)+30323 \end{aligned} \qquad (7.1.6)$$

Corectitudinea înlocuirii este demonstrată, în continuare, pentru variabila $X(k)$ (raţionamentul fiind identic şi pentru celelalte două variabile).

Fie eşantionul $X(k-1)$ obţinut cu un pas în urmă. El poate fi pus sub forma:

$$X(k-1)=177\cdot q+r,$$

unde: $q=X(k-1)/177$ (împărţire de întregi);

$r=X(k-1)\bmod 177$.

Pentru a obţine eşantionul următor, relaţia (2.1.4) cere o înmulţire,

$$\begin{aligned} 171\cdot X(k-1)&=171\cdot(177\cdot q+r)=171\cdot 177\cdot q+2\cdot q-2\cdot q+171\cdot r\\ &=30269\cdot q-2\cdot q+171\cdot r \end{aligned}$$

urmată de o operaţie modulo:

$$\begin{aligned} \left[171\cdot X(k-1)\right]\bmod 30269&=\left[171\cdot r-2\cdot q\right]\bmod 30269\\ &=\left\{171\cdot\left[X(k-1)\right]\bmod 177-2\cdot\left[X(k-1)/177\right]\right\}\bmod 30269 \end{aligned}$$

Deoarece diferenţa nu depăşeşte în valoare absolută numărul 30269 rezultă că algoritmul descris de ecuaţia (7.1.6) este corect.

Desigur că implementarea unui GNA pe baza ecuaţiei (7.1.6) va consuma mai mult timp de calcul decât în cazul utilizării ecuaţiei (7.1.4). Supus unor teste extensive, s-a dovedit că algoritmul oferă proprietăţi statistice satisfăcătoare, perioada sa fiind de aproximativ $7\cdot 10^{12}$.

Dacă dispunem de calculatoare cu viteze de calcul înalte, se poate opta pentru folosirea **algoritmului Super, Mc Laren-Marsaglia**, ce oferă proprietăţi statistice superioare. Ideea este de a combina funcţionarea a două generatoare de numere pseudoaleatorii, G1 şi G2, în următorul mod: un tablou T alcătuit din k elemente memorează primele k valori generate de G1; G2 generează un întreg j uniform distribuit pe mulţimea $\{1,2,\cdots,k\}$; se încarcă variabila de ieşire a generatorului compus, U, cu numărul aflat în locaţia $T[j]$ care apoi este modificată cu o nouă valoare generată de G1. **Algoritmul** rezultat se poate sistematiza astfel:

1. Iniţializări: folosind G1, generează k numere $X[i]$ pseudoaleatorii uniform distribuite pe $(0,1)$ şi încarcă tabloul: $T[i]=X[i]$, $1\le i\le k$

2. Generează cu G2 următorul număr pseudoaleatoriu Y uniform distribuit pe $(0,1)$
3. Calculează noul indice: $j \leftarrow \lceil k \cdot Y \rceil$, unde $\lceil \cdots \rceil$ este operaţia de rotunjire la întreg
4. $U \leftarrow T[j]$
5. Generează cu G1 un nou X
6. $T[j] = X$
7. Oferă valoarea U

Perioada secvenţei generate este $\lambda = \lambda_1 \cdot \lambda_2$, unde λ_1 şi λ_2 reprezintă perioadele generatoarelor G1 şi G2.

7.2 Metode de generare a numerelor aleatorii distribuite arbitrar

7.2.1. Metoda transformării analitice

Se bazează pe funcţia de distribuţie, $F_Z(z)$, **cdf** (***c**umulative **d**istribution **f**unction*), a variabilei aleatorii Z ce se doreşte a fi generată. În acest caz, variabila aleatorie Z se obţine utilizând relaţia:

$$Z = F_Z^{-1}(U) \qquad (7.2.1.1)$$

unde U este o variabilă aleatorie distribuită uniform pe intervalul $(0;1)$.

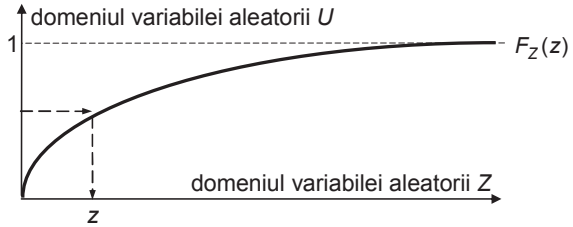


Figura 7.2.1.1: Generarea numerelor aleatorii prin metoda transformării analitice

Procedeul este prezentat sugestiv în figura 7.2.1.1, conform căreia, funcţia de distribuţie a variabilei aleatorii $Z = F_Z^{-1}(U)$ este:

$$P[Z \le z] = P\left[F_Z^{-1}(U) \le z\right] = P[U \le F_Z(z)]$$

Deoarece $P[U \le u] = u$, când U este uniform distribuit pe $(0,1)$ şi $0 \le u \le 1$, rezultă: $P[Z \le z] = F_Z(z)$, ceea ce era de dorit.

Algoritmul care implementează această metodă este:

1. Generează U uniform distribuită în $(0,1)$;
2. Calculează $Z = F_Z^{-1}(U)$;
3. Oferă valoarea Z.

Aplicaţia 7.2.1.1

Determinaţi un algoritm care, pentru a genera variabile aleatorii Bernoulli, utilizează metoda transformării analitice.

Rezolvare: Funcţia de distribuţie a unei variabile aleatorii Bernoulli, ce are probabilitatea succesului p, este reprezentată în figura 7.2.1.2. Deci:

$$Z = \begin{cases} 0 & \cdots \quad U \le p \\ 1 & \cdots \quad U > p \end{cases}$$

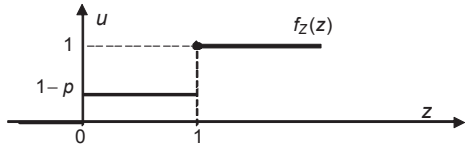


Figura 7.2.1.2: Funcţia de distribuţie a unei variabile aleatoare Bernoulli

Conform acestei relaţii algoritmul de generare a unei variabile aleatorii Bernoulli de parametru p este următorul:

1. Generează U uniform distribuită în $(0,1)$;
2. Dacă $U \le 1-p$: $Z = 0$;
3. Dacă $U > 1-p$: $Z = 1$;
4. Oferă Z.

* * *

Aplicaţia 7.2.1.2 **(distribuţia exponenţială)**

Determinaţi un algoritm care, pentru a genera variabile aleatorii distribuite exponenţial, utilizează metoda transformării analitice.

Rezolvare: Întrucât: $F_Z(z) = 1 - \exp(-\lambda z)$ avem:

$$U = 1 - \exp(-\lambda Z) \text{ sau } Z = (-1/\lambda) \cdot \ln(1-U) \qquad (7.2.1.2)$$

Deoarece atât U cât şi $1-U$ sunt uniform distribuite în intervalul (0,1) putem evita o scădere scriind:

$$Z = (-1/\lambda) \cdot \ln(U) \qquad (7.2.1.3)$$

Desigur că, secvenţele generate de cele două formule nu vor fi identice, dar vor fi echivalente din punct de vedere statistic.

Conform relaţiei (7.2.1.3), algoritmul de generare a unei variabile aleatorii exponenţiale de parametru λ urmează secvenţa descrisă în continuare:

1. Generează U uniform distribuită în $(0,1)$;

2. Calculează $Z=(-1/\lambda)\cdot\ln(U)$;
3. Oferă valoarea Z.

Aplicaţia 7.2.1.3 **(distribuţia geometrică)**

Determinaţi un algoritm care, pentru a genera variabile aleatorii distribuite geometric, utilizează metoda transformării analitice.

Rezolvare: Fie X o variabilă aleatorie geometrică descrisă de probabilităţile elementare: $P[X=k]=p\cdot(1-p)^{k-1}$, unde:

$k=$ numărul de repetări până apare primul succes;
$p=$ probabilitatea succesului unei încercări.

Fie Z o variabilă aleatorie exponenţială pentru care avem:

$$P[k<Z\le k+1]=1-\mathrm{e}^{-\lambda(k+1)}-1+\mathrm{e}^{-\lambda k}=\mathrm{e}^{-\lambda k}\cdot(1-\mathrm{e}^{-\lambda}).$$

Dacă se consideră $p=1-\mathrm{e}^{-\lambda}$, rezultă că:

$$P[k<Z\le k+1]=p\cdot(1-p)^k=P[X=k+1].$$

Deci, obţinem o variabilă aleatorie geometrică, X, de parametru p prin generarea unei variabile aleatorii, Z, exponenţiale, de parametru $\lambda=-\ln(1-p)$ şi majorarea sa la întreg. Rezultă următorul algoritm de generare de variabile aleatorii distribuite geometric de parametru p:

1. Generează U;
2. Calculează $Z=\dfrac{\ln U}{\ln(1-p)}$;
3. Calculează $X=\lceil Z\rceil$, unde $\lceil*\rceil$ reprezintă operaţia de majorare la întreg;
4. Oferă X.

* * *

Aplicaţia 7.2.1.4 (distribuţii descrise prin histograme)

Determinaţi un algoritm care aplică metoda transformării analitice în cazul în care funcţia de distribuţie şi/sau inversa ei nu au expresii explicite de calcul.

Rezolvare:

În acest caz, pe baza funcţiei densitate de probabilitate, se construieşte histograma repartiţiei (vezi cap. 5.3 Histograma), cu ajutorul căreia se determină funcţia de distribuţie aproximativă, în trepte:

$$F_i=\sum_{j=1}^{i}p_j\text{, cu } i=\overline{0,N},$$

unde $p_1,\cdots,p_N$ sunt probabilităţile celor N intervale din figura 7.2.1.3.

Se obţine astfel următorul algoritm:

1. Generează U uniform distribuită pe $(0,1)$;
2. Găseşte $i=\overline{1,N}$ pentru care $F_{i-1}<U\le F_i$;

3. Calculează $Z=Z_{i-1}+\dfrac{U-F_{i-1}}{c_i}$;
7. Oferă Z.

Observaţie: La pasul 3, variabila aleatorie Z se obţine printr-o interpolare pe intervalul $[z_{i-1},z_i]$.

* * *

Aplicaţia 7.2.1.5 **(distribuţii discrete)**

Determinaţi un algoritm care să utilizeze metoda transformării analitice pentru a genera variabile aleatorii discrete, cu spaţii de eşantionare:
i) finite; ii) infinit numărabile.

Rezolvare: i) Variabila aleatorie Z ia valorile $z_1,z_2,\cdots$ cu probabilitatea $P[Z=z_k]=p_k$. Funcţia sa de distribuţie este: $F_k=\sum_{j=1}^{k}p_j$. Deci, algoritmul pentru generarea variabilelor aleatorii discrete cu spaţiu de eşantionare finit este:

1. $k=1$;
2. Generează U uniform distribuită pe $(0,1)$;
3. Dacă $U\le F_k$: $Z=z_k$, sări la 5;
4. $k=k+1$, sări la 3;
5. Oferă Z.

ii) În acest caz nu se pot calcula aprioric valorile F_k pentru toate realizările z_k ale variabilei aleatorii Z, pentru a fi numerotate şi ulterior utilizate. De aceea, calculul valorilor F_k se reia ***complet*** de la un apel la altul al generatorului sau ***parţial***, valorile funcţiei de distribuţie ce corespund realizărilor frecvente memorându-se aprioric într-un tablou. În ambele situaţii este necesar ca procedeul de calcul al probabilităţilor să fie cât mai eficient. De exemplu, pentru distribuţia Poisson se utilizează relaţia de recurenţă: $p_{k+1}/p_k=\alpha/(k+1)$.

Pe baza acestor observaţii se poate schiţa următorul algoritm cu reluare completă a calculului valorilor F_k pentru generarea variabilelor aleatorii discrete, cu spaţiu de eşantionare infinit numărabil:

1. $C=p_0$, $B=C$, $k=0$;
2. Generează U uniform distribuită pe $(0,1)$;
3. Dacă $U\le B$: $Z=z_k$, treci la 5;
4. $C=C\cdot(p_{k+1}/p_k)$, $B=B+C$ şi $k=k+1$, treci la 3;
5. Oferă Z.

* * *

Aplicaţia 7.2.1.6

Determinaţi numărul mediu de paşi necesari algoritmilor de generare a

variabilelor aleatorii discrete atunci când căutarea începe de la începutul spaţiului de eşantionare. Cazuri particulare:

i) distribuţie binomială cu $n=5$ şi $p=1/2$; ii) distribuţie Poisson; ş.a.m.d..

Rezolvare: Fie N o variabilă aleatorie ce reprezintă numărul de paşi. În acest caz, media sa este dată de relaţia:

$$\begin{aligned}\mathrm{E}[N] &= 1\cdot P[U\le p_0]+2\cdot P[p_0<U\le p_0+p_1]+\ldots\\ &\quad +(n+1)\cdot P[p_0+\ldots+p_{n-1}<U\le p_0+\ldots+p_n]\\ &= 1\cdot p_0+2\cdot p_1+\ldots+(n+1)\cdot p_n\end{aligned}$$

Adică: $\mathrm{E}[N]=\mathrm{E}[Z]+1$, unde Z este variabila aleatorie ce se doreşte a se genera. Pentru cazul particular: i) $\mathrm{E}[N]=2{,}5+1=3{,}5$.

* * *

Aplicaţia 7.2.1.7

Determinaţi numărul mediu de paşi necesari algoritmilor de generare a variabilelor aleatorii discrete atunci când căutarea începe din interiorul spaţiului de eşantionare. Caz particular: distribuţie binomială cu $n=5$, $p=1/2$.

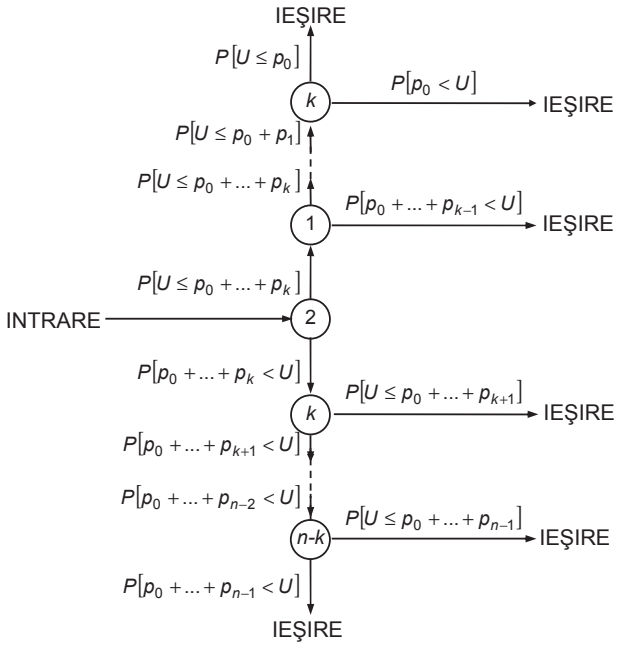


Figura 7.2.1.4: Diagrama numărului de căutări ce pornesc din interiorul spaţiului de eşantionare

Rezolvare: În general, acest procedeu se aplică pentru variabile aleatorii discrete cu spaţiu de eşantionare finit. În acest caz, procesul de căutare poate urma diagrama din figura 7.2.1.4 în ale cărei noduri este trecut numărul căutărilor până în acel punct. Conform acestei figuri, media numărului de căutări, N, este:

$$\begin{aligned}\mathrm{E}[N] &= P\left[U\le\sum_{j=0}^{k}p_j\right]\cdot\sum_{j=2}^{k}j\cdot P\left[\sum_{s=0}^{k-j+1}p_s<U\right]\\ &\quad +P\left[\sum_{j=0}^{k}p_j<U\right]\cdot\sum_{j=2}^{n-k}j\cdot P\left[U\le\sum_{s=0}^{k+j-1}p_s\right]\\ &= \sum_{j=2}^{k}j\cdot\sum_{s=k}^{k+2-j}p_s+\sum_{j=2}^{n-k}j\cdot\sum_{s=k+1}^{k-j+3}p_s\end{aligned}$$

Pentru cazul particular avem: $\mathrm{E}[N]=3\cdot p_1+5\cdot p_2+5\cdot p_3+3\cdot p_4$

cu: $p_1=p_4=\mathrm{C}_5^1\cdot(1/2)^5=5/32$ şi $p_2=p_3=\mathrm{C}_5^2\cdot(1/2)^5=10/32$.

În final se obţine că: $\mathrm{E}[N]=4{,}06$.

După cum era de aşteptat, valoarea obţinută este mai mică decât în cazul exemplului anterior, diferenţa accentuându-se cu cât spaţiul de eşantionare conţine mai multe realizări.

7.2.2. Metoda respingerii

Se bazează pe funcţia densitate de probabilitate $f_Z(z)$, **pdf** (***p**robability **d**ensity **f**unction*), a variabilei aleatorie Z ce se doreşte a se genera. Pentru a o explica, vom presupune iniţial că:

- $f_Z(z)$ este nulă în afara intervalului $[0,a]$;
- $f_Z(z)$ ia valori în intervalul $[0,b]$.

În această situaţie, metoda funcţionează astfel:

1 - Generează X distribuit uniform în $(0,a)$;

2 - Generează Y distribuit uniform în $(0,b)$;

3 - Dacă $Y\le f_Z(X)$: acceptă X, deci $Z=X$, sări la pasul 5;

4 - Respinge X şi revine la pasul 1;

5 - Oferă Z.

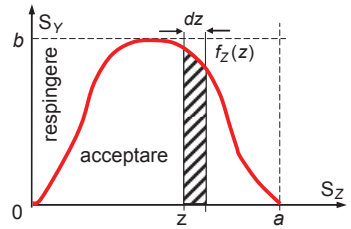


Figura 7.2.2.1: Metoda respingerii când funcţia $f_Z(z)$ este mărginită şi are domeniul finit

Conform figurii 7.2.2.1, primii trei paşi împart careul de arie $a \cdot b$ în două regiuni despărţite de curba funcţiei densitate de probabilitate. Probabilitatea acceptării este raportul dintre aria de sub curba funcţiei densitate de probabilitate şi aria careului, deci: $P\left[X_{\text{acceptat}}\right] = 1/(a \cdot b)$. În acest caz:

$$P\left[z < X \le z + \mathrm{d}z \,\middle|\, X_{\text{acceptat}}\right] = \frac{P\left[\{z < X \le z + \mathrm{d}z\} \,\&\, \{X_{\text{acceptat}}\}\right]}{P\left[X_{\text{acceptat}}\right]}$$

$$= \frac{\text{aria haşurată}/(a \cdot b)}{1/(a \cdot b)} = f_Z(z)$$

Deci, când X este acceptat, atât el, cât şi variabila aleatorie de ieşire, Z, respectă distribuţia dorită. Metoda, aşa cum a fost descrisă mai sus, prezintă două inconveniente. Primul este generat de faptul că în multe situaţii, zona de respingere este atât de mare încât numărul de repetări până la apariţia unei acceptări poate să devină excesiv de mare. Al doilea este datorat restricţiilor impuse funcţiei densitate de probabilitate ce nu permit utilizarea sa în cazul în care $f_Z(z)$ este nemărginită şi/sau de domeniu infinit. Efectul primului neajuns este diminuat, iar a doua deficienţă este eliminată de versiunea generală a metodei respingerii, prezentată în continuare.

Metoda generală de respingere presupune stabilirea unei variabile aleatorii, W, descrisă prin $f_W(z)$ care este *uşor* de generat şi care pentru o constantă $K > 1$ îndeplineşte, pentru orice z, relaţia: $K \cdot f_W(z) \ge f_Z(z)$.

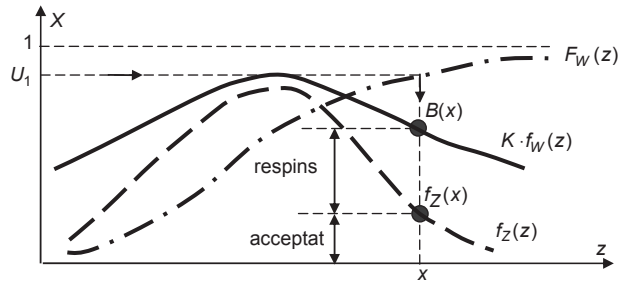


Figura 7.2.2.2: Metoda generală a respingerii

În acest fel, conform figurii 7.2.2.2, regiunea de sub $K \cdot f_W(z)$ conţine curba $f_Z(z)$, iar metoda respingerii se desfăşoară după cum urmează:

1. Generează X cu distribuţie $f_W(x)$ (se utilizează metoda transformării analitice);
2. Stabileşte intervalul pe care se desfăşoară testul metodei: $B = K \cdot f_W(x)$;
3. Generează Y uniform distribuit în $(0, B)$;

4. Dacă $Y \le f_Z(X)$: acceptă X, deci $Z = X$, sări la pasul 6;
5. X este respins, sari la pasul 1;
6. Oferă Z.

Faptul că variabila aleatorie, Z, urmează distribuţia impusă se demonstrează ca în cazul precedent, diferenţa constând în noua suprafaţă ce cuprinde cele două regiuni şi care este de arie K. Eficienţa de calcul pentru acest algoritm depinde de cât de "aproape" este curba $K \cdot f_W(z)$ de curba $f_Z(z)$.

Aplicaţia 7.2.2.1 **(distribuţia gama, $0 < \alpha < 1$)**

Să se aplice metoda respingerii pentru a genera o variabilă aleatorie, Z, ce urmează o distribuţie gama de parametri $0 < \alpha < 1$ şi $\lambda = 1$.

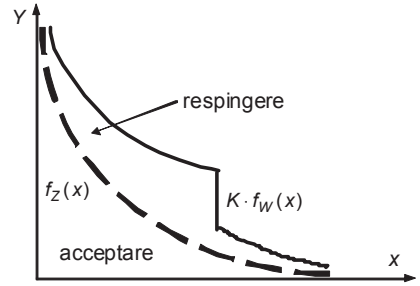


Figura 7.2.2.3: Metoda respingerii în cazul distribuţiei gama cu $0 < \alpha < 1$ şi $\lambda = 1$

Rezolvare: Conform figurii 7.2.2.3, funcţia $K \cdot f_W(x)$ ce "acoperă" $f_Z(x)$ este:

$$f_Z(x) = \frac{x^{\alpha-1} \cdot \mathrm{e}^{-x}}{\Gamma(\alpha)} \le K \cdot f_W(x) = \begin{cases} x^{\alpha-1}/\Gamma(\alpha) & \text{pentru } 0 \le x \le 1 \\ \mathrm{e}^{-x}/\Gamma(\alpha) & \text{pentru } x > 1 \end{cases}$$

Funcţia densitate de probabilitate, $f_W(x)$, ce corespunde termenului drept este:

$$f_W(x) = \begin{cases} \alpha \cdot \mathrm{e} \cdot x^{\alpha-1}/(\alpha + \mathrm{e}) & \text{pentru } 0 \le x \le 1 \\ \alpha \cdot \mathrm{e} \cdot \mathrm{e}^{-x}/(\alpha + \mathrm{e}) & \text{pentru } x > 1 \end{cases}$$

Integrând, se obţine: $F_W(x) = \begin{cases} \mathrm{e} \cdot x^{\alpha}/(\alpha + \mathrm{e}) & \text{pentru } 0 \le x \le 1 \\ 1 - \alpha \cdot \mathrm{e}^{1-x}/(\alpha + \mathrm{e}) & \text{pentru } x > 1 \end{cases}$

Această funcţie se inversează, pentru a putea aplica metoda transformării analitice, în cadrul primului pas al metodei de respingere:

$$F_W^{-1}(u) = \begin{cases} \left[(\alpha + \mathrm{e}) \cdot u/\mathrm{e}\right]^{1/\alpha} & \text{pentru } u \le \mathrm{e}/(\alpha + \mathrm{e}) \\ -\ln\left[(\alpha + \mathrm{e}) \cdot (1-u)/(\alpha \cdot \mathrm{e})\right] & \text{pentru } u > \mathrm{e}/(\alpha + \mathrm{e}) \end{cases}$$

Valoarea constantei K, necesară pasului următor, este: $K = \frac{\alpha + e}{\alpha \cdot e \cdot \Gamma(\alpha)}$

Observaţie: Dacă valoarea acceptată o amplificăm cu λ obţinem o variabilă aleatorie gama de parametri $0 < \alpha < 1$ şi λ.

* * *

Aplicaţia 7.2.2.2

Justificaţi expresia de calcul a densităţii de probabilitate precizată în aplicaţia precedentă.

Indicaţie: A se vedea modul prin care orice funcţie pozitivă şi continuă pe intervale poate deveni o densitate de probabilitate.

* * *

Aplicaţia 7.2.2.3 **(distribuţia gama, $\alpha > 1$)**

Ştiind că $f_W(x) = \frac{\lambda \alpha^{\lambda} \cdot x^{\lambda-1}}{(\alpha^{\lambda} + x^{\lambda})^2}$ şi $K = \sqrt{2\alpha - 1}$ oferă o funcţie $K \cdot f_W(x)$ ce mărgineşte funcţia densitate de probabilitate a unei variabile aleatorii gama cu $\alpha > 1$, să se determine $F_W(x)$.

Rezulzat: $F_W(x) = \frac{x^{\lambda}}{\alpha^{\lambda} + x^{\lambda}}$, pentru $x > 0$.

7.2.3. Generarea funcţiilor de variabile aleatorii

Având la dispoziţie metode simple de generare a variabilei aleatorii X se poate obţine variabila aleatorie definită prin $Y = g(X)$. Desigur, afirmaţia se poate extinde şi în cazul $Z = h(X_1, X_2, \cdots, X_n)$ unde $X_1, X_2, \cdots, X_n$ sunt ieşirile unor GNA corespunzătoare.

Aplicaţia 7.2.3.1 **(distribuţia m-Erlang)**

Determinaţi un algoritm care nu utilizează metoda transformării analitice pentru a genera variabile aleatorii ce urmează o distribuţie *m*-Erlang.

Rezolvare: Deoarece o variabilă aleatorie *m*-Erlang se obţine prin adunarea a m variabile aleatorii distribuite exponenţial identice şi independente se deduce următorul algoritm:

1. $X = 0$; $C = m$;
2. Generează o variabilă aleatorie exponenţială, Z, de parametru λ;
3. $X = X + Z$;
4. Dacă $C = 1$: sari la pasul 6;
5. $C = C - 1$; revino la pasul 2;
6. Oferă X.

Observaţie: Pentru valori ale lui m mari este de preferat utilizarea metodei respingeri aplicată pentru o variabilă aleatorie gama, de parametri $\alpha = m$ (ce devine o variabilă aleatorie *m*-Erlang).

* * *

Aplicaţia 7.2.3.2 **(distribuţia Poisson)**

Determinaţi un algoritm care generează o variabilă aleatorie Poisson de parametru α.

Rezolvare: O variabilă aleatorie Poisson, X, reprezintă numărul de evenimente ce au loc într-un anumit interval, ΔT. Dar timpii Z_i dintre două sosiri succesive sunt variabile aleatorii independente identic distribuite exponenţial de parametru $\lambda = \alpha / \Delta T$. De aceea o variabilă aleatorie Poisson reprezintă numărul de variabile aleatorii Z_i care însumate nu depăşesc intervalul ΔT.

Cum variabila aleatorie exponenţială se obţine prin intermediul variabilei aleatorii uniforme se poate scrie relaţia:

$$Z_0 + Z_1 + \ldots = -\frac{\ln U_0 + \ln U_1 + \ln U_2 + \ldots}{\lambda},$$

care, la fiecare adăugare de termen, trebuie comparată cu ΔT pentru a sesiza pasul la care această relaţie este depăşită:

$$-\frac{\ln(U_0 \cdot U_1 \cdot \ldots)}{\lambda} > \Delta T.$$

Cum $\lambda = \alpha / \Delta T$ rezultă că trebuie verificată inegalitatea:

$$U_0 \cdot U_1 \cdot U_2 \cdot \ldots < e^{-\alpha} \qquad (7.2.3.1)$$

Algoritmul corespunzător, de generare a unei variabile aleatorii Poisson de parametru α, are următoarea structură:

1. $A = 1$ şi $k = 0$; (k = numărul de apariţii);
2. Generează U_k distribuit uniform pe $(0,1)$;
3. $A = A \cdot U_k$;
4. Dacă $A < e^{-\alpha}$: $X = k$, sări la 6;
5. $k = k + 1$, sări la 2;
6. Oferă X.

* * *

Aplicaţia 7.2.3.3 **(distribuţia Gauss)**

Să se determine o metodă de generare a unei variabile aleatorii Gauss de medie 0 şi varianţă 1.

Rezolvare:

Soluţia 1: Teorema limită centrală arată că o sumă de variabile aleatorii independente identic distribuite de medie finită μ şi varianţă finită σ^2 tinde să aibă o distribuţie Gauss de medie $n \cdot \mu$ şi varianţă $n \cdot \sigma^2$. Din acest motiv, soluţia cea mai simplă de generare a unei variabile aleatorii Gauss de medie $\mu_r = 0$ şi varianţă $\sigma_r^2 = 1$ este dată de formula:

$$Y = \sum_{k=1}^{n} U(k) - n/2 \qquad (7.2.3.2)$$

în care $U(k)$, cu $k=\overline{1,n}$, este un set de variabile aleatorii independente, identic distribuite pe $(0,1)$ după o lege uniformă.

Este de remarcat că această formulă limitează domeniul $(-\infty,+\infty)$ al unei variabile aleatorii Gauss la intervalul $[-n/2,n/2]$. Pentru a realiza un compromis între acurateţe şi viteza de calcul, în practică se utilizează formula (7.2.3.2) în cazul $n=12$ (dar, bineînţeles, se poate merge şi mai departe).

Soluţia 2: Conform aplicaţiei 3.6.3, transformarea în coordonate polare a coordonatelor carteziene, generate de două variabile aleatorii Gauss i.i.d., de medie 0 şi varianţă 1, produce două variabile, de asemenea independente, una uniformă, a unghiului, şi cealaltă, a razei, cu distribuţie Rayleigh.

Mai mult, aplicaţia 2.3.1 a arătat că pătratul unei variabile Rayleigh de parametru σ^2 este o variabilă exponenţială de parametru $(2\sigma^2)^{-1}$. Prin urmare, plecând de la relaţiile de transformare inversă:

$$X=R\cdot\cos\Theta \text{ şi } Y=R\cdot\sin\Theta \qquad (7.2.3.3)$$

se poate implementa următorul algoritm de generare a două variabile aleatorii Gauss independente de medie 0 şi varianţă 1 (**Metoda Box-Muller**):

1. Generează U_1 şi U_2 variabile aleatorii independente şi uniform distribuite în $(0,1)$;
2. $R^2=-2\cdot\ln U_1$ (generarea unei variabile aleatorii exponenţiale de parametru $0{,}5$);
3. $\Theta=2\pi\cdot U_2$;
4. $X=R\cdot\cos\Theta$ şi $Y=R\cdot\sin\Theta$, conform relaţiei (7.2.3.3);
5. Oferă X şi Y.

Observaţii:

- Teoretic, această metodă produce numere pe întreg intervalul: $(-\infty,+\infty)$. Luând în considerare însă precizia oferită de calculator, algoritmul va genera numere în intervalul $\left(-\sqrt{2\ln M},+\sqrt{2\ln M}\right)$, unde M reprezintă cel mai mare întreg ce poate fi reprezentat în memoria calculatorului (numărul minim produs de un GNA distribuite uniform pe $(0,1)$ este $1/M$, şi acesta calculat cu o anumită precizie).
- De asemenea, viteza de calcul este mult scăzută faţă de prima soluţie. Aceasta se datorează complexităţii sporite a operaţiilor matematice ce le implică ($\ln$, $\cos$, $\sin$, $\sqrt{\ }$, $*$).

* * *

Aplicaţia 7.2.3.4 **(distribuţia hiperexponenţială; metoda amestecării)**

Să se determine un algoritm de generare a unei variabila aleatorii

hiperexponenţiale în două trepte. Să se generalizeze rezultatul.

Rezolvare: Funcţia densitate de probabilitate a unei astfel de variabile aleatorii este:

$$f_X(x)=p\,a\cdot\exp(-a\cdot x)+(1-p)\,b\cdot\exp(-b\cdot x) \qquad (7.2.3.4)$$

Expresia (7.2.3.4) arată că variabila aleatorie X este un "amestec" (o compunere) a două variabile aleatorii, X_1 şi X_2, exponenţiale, de parametru a, respectiv b, controlat de o variabila aleatorie discretă, N, de tip Bernoulli. Pentru a genera X este necesar ca iniţial să se realizeze un experiment Bernoulli şi, în funcţie de rezultatul acestuia, să se genereze ori X_1, ori X_2.

În cazul general: $f_X(x)=\sum_{i=1}^{n}p_i\cdot f_{X_i}(x)$, în care $\{p_i\}$ reprezintă probabilităţi de repartiţie ale unei variabile aleatorii N. Conform relaţiei anterioare, rezultă următorul algoritm de amestecare discretă:

1. Se generează N (utilizând, de exemplu, algoritmul descris în *aplicaţia 7.2.1.5*);
2. $i=N$;
3. Generează X_i; $Z=X_i$;
4. Oferă Z.

capitolul 8

Teste statistice

Aprecieri relative la proprietăţile statistice ale unui fenomen aleatoriu pot fi obţinute prin supunerea acestuia la diverse teste. Cu ajutorul lor se stabilesc atât proprietăţile temporare (staţionaritatea, autocorelaţia etc.) cât şi proprietăţi pe ansamblu (media, varianţa, funcţia de repartiţie etc.) ale fenomenului studiat.

În cazul generatoare de numere aleatorii, GNA, aceste teste se aplică în special generatoarelor de numere uniform distribuite în $(0,1)$, ele fiind cele care dictează, în principal, buna funcţionare a celorlalte.

8.1 Testul de staţionaritate

Noţiunea de staţionaritate este descrisă la modul riguros în capitolul următor, dedicat proceselor aleatorii. La modul intuitiv, suficient expunerii prezente, ea reprezintă o măsură a ***regularităţii statistice***, de care poate da dovadă un anumit fenomen aleatoriu, în sensul că medii, respectiv varianţe, pe eşantioane diferite sunt apropiate ca valoare. În consecinţă, verificarea staţionarităţii, luând în considerare aceste două mărimi, poate urma procedură de mai jos:

- Se generează o secvenţă lungă de n realizări.
- Se împarte secvenţa în m segmente disjuncte şi se calculează media şi varianţa pe fiecare segment (vezi relaţia 6.3.3.1, respectiv 6.4.2.2).
- Se verifică dacă mediile şi varianţele se găsesc în vecinătăţi acceptabile din punct de vedere statistic.

8.2 Testul de autocorelaţie

Evaluarea autocorelaţiei ce caracterizează o secvenţă de numere aleatorii este extrem de utilă în cazul GNA care trebuie să ofere secvenţe

necorelate şi de perioadă suficient de mare pentru a nu se relua pe durata unei simulări. Dacă testul de staţionaritate este îndeplinit, atunci coeficientul de autocorelaţie este dat de expresia:

$$\rho_{XX}(k)=\frac{\mathrm{E}[X(j)\cdot X(j+k)]-\mu_x^2}{\mathrm{E}\left[X^2(j)\right]-\mu_X^2}\overset{\text{notat}}{=}\frac{\mathrm{COV}_{XX}(k)}{\mathrm{COV}_{XX}(0)} \qquad (8.2.1)$$

În cazul ideal:

$$\rho_{XX}(k)=\begin{cases}1 & \text{pentru } k=0\\ 0 & \text{pentru } k\neq 0\end{cases} \qquad (8.2.2)$$

Ansamblul coeficienţilor de autocorelaţie, ordonaţi după **decalajul** (**lag**) k, se numeşte **funcţia de autocorelaţie**. Atunci când $\rho_{XX}(k)\neq 0$, pentru $k\neq 0$, înseamnă că secvenţa generată este corelată, iar dacă $\rho_{XX}(k)$ are "vârfuri" la $k=m\cdot n_0$ şi $m=1,2,\cdots$, ea este şi periodică, de perioadă n_0.

Deoarece verificarea practică a corelaţiei şi a periodicităţii utilizează **estimatorul funcţiei de autocorelaţie**, $\hat{\rho}_{XX}(k)$, este necesar ca interpretarea rezultatelor oferite de test să ţină cont de variabilitatea statistică a acestui estimator.

În practică se poate utiliza următoarea procedură de testare a corelaţiei şi a periodicităţii:

- Se generează o secvenţă foarte lungă, de lungime L;
- Se împarte secvenţa în **grupe** (***batch***) a câte n realizări ce se suprapun în $n/2$ poziţii (vezi figura 8.2.1 în care **pasul** (**step**), s, este egal cu $n/2$). Se obţin astfel vectorii $\mathbf{X}_0,\mathbf{X}_1,\cdots,\mathbf{X}_k,\cdots$;
- Se calculează coeficienţii de corelaţie între $\mathbf{X}_0$ şi $\mathbf{X}_0,\mathbf{X}_1,\cdots$ folosind relaţia:

$$\hat{\rho}_{XX}(k\cdot s)=\hat{\mathrm{COV}}(k\cdot s)\big/\hat{\mathrm{COV}}(0) \qquad (8.2.3)$$

unde:

$$\hat{\mathrm{COV}}(k\cdot s)=\frac{1}{n-1}\sum_{i=1}^{n}X_0(i)\,X_k(i)-m_{\mathbf{X}_0}\cdot m_{\mathbf{X}_k} \qquad (8.2.4)$$

Reducerea numărului de operaţii se face considerând un pas, s, cât mai mare, dar nu peste $n/2$ (uzual, $s=n/2$, ca în figura 8.2.1) şi utilizând tehnica transformatei Fourier rapide, **FFT** (***F**ast **F**ourier **T**ransform*), sau programele specializate din mediul de calcul MatLab;

- Se trasează $\left|\hat{\rho}_{X_1X_k}(k\cdot s)\right|$ pentru fiecare k şi se caută vârfuri (vezi figura 8.2.1). Precizia cu care se determină perioada, n_0, a generatorului se controlează prin mărimea valorii s care, iniţial, se alege $s=n/2$, iar apoi, pentru zona de maxim, căutarea se face cu paşi mai mici, de exemplu $s=n/10$.

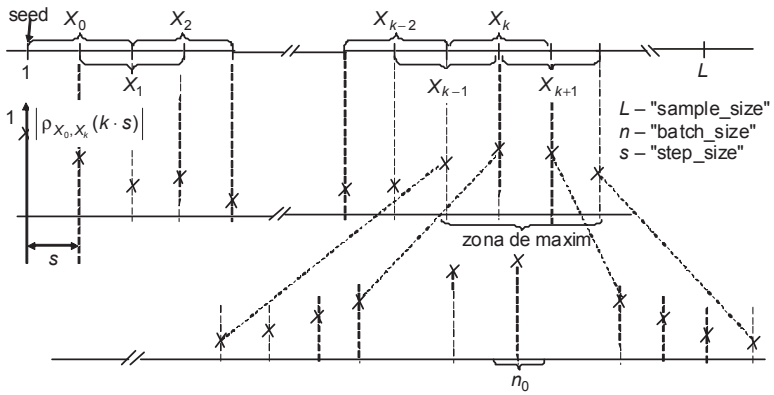


Figura 8.2.1: Principiul testului de autocorelaţie

În general, testarea GNA pentru periodicitate cere un volum de calcul substanţial. Aceasta se datorează evaluării repetate a expresiei $\hat{\rho}_{X\,X}(k)$ pe segmente de dimensiuni mari atunci când perioada căutată este foarte mare şi reluării testului pentru diverse valori ale punctului din care este reprodusă secvenţa. Acele valori ce conduc la perioade mari trebuie reţinute pentru a fi utilizate ulterior la simulările de lucru, în care GNA controlează diverse fenomene (precum generări de pachete, timpi de servire etc.).

8.3 Histograma

Histograma (coloanele albe din figurile 8.3.1 şi 8.3.2) este reprezentarea grafică a frecvenţelor absolute, adică a numărului de apariţie pe parcursul unui număr de repetări, ale unor evenimente legate de un anumit experiment.

Astfel, luând drept exemple fenomene aleatorii, care implică sistemele de comutaţie, putem considera că histograma din figura 8.3.1 s-a trasat în urma recepţionării din liniile de abonaţi a 114 cifre, iar histograma din figura 8.3.2 s-a obţinut în urma observării a 720 de convorbiri telefonice, evenimentele urmărite fiind:

- în primul caz, de genul: $\{\text{din linie s-a recepţionat cifra } i\}$,
- în al doilea caz: $\{\text{durata convorbirii se încadrează în intervalul } [a,b]\}$

Histograma reprezintă o metodă uzuală de interpretare grafică a datelor experimentale, care constituie deseori o etapă intermediară pe parcursul desfăşurării unei analize statistice, în vederea identificării distribuţiei teoretice care descrie cât mai precis fenomenul studiat. În acest caz, pasul următor constă în alegerea unei anumite distribuţii şi trasarea, alături de coloanele histogramei, a coloanelor (negre) corespunzătoare numărului de apariţii de aşteptat în condiţiile în care experimentul s-ar desfăşura conform acestor distribuţii. Astfel, prin

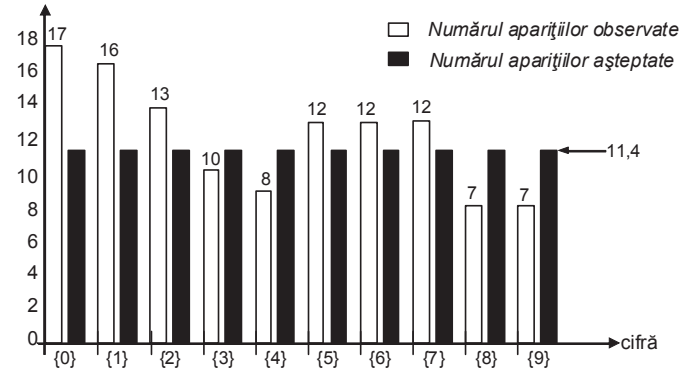


Figura 8.3.1: Histogramă pentru compararea datelor experimentale cu cele teoretice - cazul distribuţiei uniforme

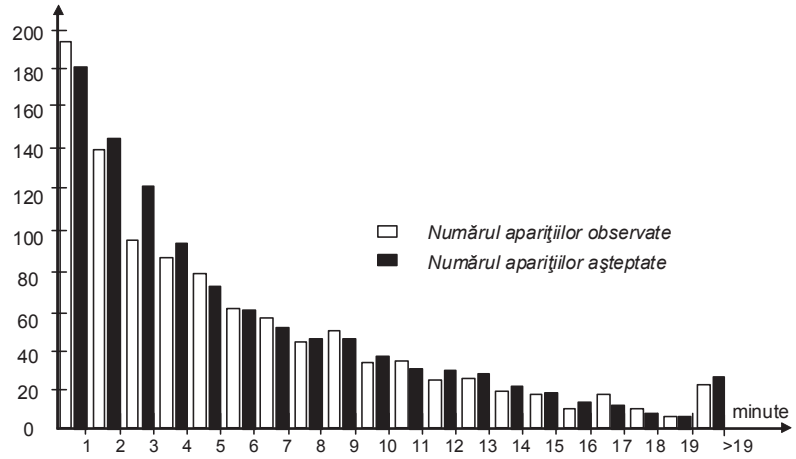


Figura 8.3.2: Histogramă pentru compararea datelor experimentale cu cele teoretice - cazul distribuţiei exponenţiale

alăturarea celor două seturi de coloane, este posibilă compararea vizuală pe baza căreia se poate decide dacă ipoteza (distribuţia) considerată este acceptabilă. Dat fiind natura subiectivă a comparaţiei vizuale, aceasta este, în general, urmată, atunci când *diferenţele* nu mai sunt atât de evidente pentru a permite respingerea ipotezei, de teste de natură calitativă, în care decizia se ia în urma efectuării unor calcule matematice. Astfel de teste sunt: testul χ^2 şi testul Kolmogorov - Smirnov, ale căror principii şi mod de utilizare sunt descrise în cele ce urmează.

8.4 Testul χ^2

Testul χ^2 este o metodă cantitativă, pe baza căreia, în urma efectuării unor calcule, se decide acceptarea sau respingerea ipotezei presupuse a caracteriza fenomenul aleatoriu studiat. Pentru înţelegerea lui, considerăm un caz simplu: recepţionarea unui simbol binar dintr-o linie numerică (sau, mai puţin academic, aruncarea cu banul).

Facem următoarea ipoteză, pe care o notăm H_0: *"valoarea simbolului recepţionat urmează o lege de tip Bernoulli, cu* $p = 1/2$ *"* . Pentru a o verifica, repetăm experimentul de mai multe ori. Astfel, considerând elementele:

- n este numărul total de simboluri recepţionate,
- N este numărul de simboluri "1" recepţionate (observate),
- m este numărul aşteptat de simboluri "1", conform ipotezei H_0

se scrie următoarea egalitate:

$$P\left[\,|N-m| \geq t_\alpha \,|H_0\right] = \alpha,$$

în care: - $|N-m|$ = diferenţa dintre ce se observă şi ce se aşteaptă,

- t_α = ***deviaţia critică***;

- α = ***nivel****ul (****prag****ul) probabilistic* ***semnificativ***, ce reprezintă o valoare impusă, de obicei 1% sau 5%, a probabilităţii ca, în ipoteza H_0, diferenţa $|N-m|$ să depăşească valoarea t_α.

Decizia se ia astfel:

- H_0 este adevărată dacă, la nivelul probabilistic semnificativ impus, deviaţia critică nu este depăşită de diferenţa $|N-m|$;
- H_0 este falsă în caz contrar, adică atunci când ceea ce s-a observat s-a abătut nejustificat de mult din punct de vedere probabilistic de ceea ce era de aşteptat.

Aplicaţia 8.4.1

Considerând un nivel probabilistic semnificativ, α, de 1%, să se verifice dacă ipoteza $H_0 = \text{"biţii generaţi de o sursă sunt echiprobabili"}$ este adevărată ştiind că din 100 de biţi observaţi doar 62 au avut valoarea "1".

Rezolvare: Pe baza ipotezei, numărul aşteptat este $m = 100 \cdot p = 50$. Deviaţia critică se determină rezolvând ecuaţia:

$$P\left[\,|N-50| \geq t_\alpha \,|H_0\right] = 1 - \sum_{k=50-t_\alpha}^{50+t_\alpha} C_{100}^k \cdot (1/2)^{100} = 0{,}01$$

a cărei soluţie este: $t_\alpha = 13$. Diferenţa dintre ce se observă şi ce se aşteaptă are valoarea $|62-50| = 12$. Prin urmare, deoarece $12 < 13$, înseamnă că ipoteza avută în vedere este acceptată.

Aplicaţia 8.4.2

Reluaţi aplicaţia anterioară dacă nivelul de probabilitate semnificativ, α, este de 5%.

Indicaţie: se realizează un program în MatLab care rezolvă o ecuaţie similară celei considerate în aplicaţia precedentă.

* * *

În cazul general, testul χ^2 presupune următorii paşi:

- Se împarte spaţiul de eşantionare, S_X, în k mulţimi disjuncte;
- Se calculează, luând în considerare ipoteza H_0, asumată, probabilitatea aşteptată p_j ca rezultatul unui experiment să fie în submulţimea j, cu $j = \overline{1,k}$;
- Se fixează numărul de repetări ale experimentului n;
- Se determină numărul aşteptat de încercări ce cad în submulţimea j, notat şi dat de expresia:

$$m_j = n \cdot p_j \qquad (8.4.1)$$

- Se efectuează experimentul de n ori, determinând numărul repetărilor ce cad în submulţimile alese, N_j cu $j = \overline{1,k}$;
- Se calculează suma ponderată a pătratelor diferenţelor dintre ceea ce se observă şi ceea ce se aşteaptă:

$$D^2 = \sum_{j=1}^{k} (N_j - m_j)^2 / m_j \qquad (8.4.2)$$

Pentru un număr de repetări suficient de mare, variabila aleatorie D^2 are o funcţie densitate de probabilitate ce se poate aproxima cu cea a variabilei aleatorii χ^2 cu $k-1$ grade de libertate. În acest caz, deviaţia critică, la un nivel probabilistic semnificativ dat, rezultă din ecuaţia:

$$P\left[\,D^2 \geq t_\alpha \,|\, H_0\right] = \alpha \qquad (8.4.3)$$

unde α corespunde ariei haşurată din figura 8.4.1.

- Pentru numărul de submulţimi k, fixat anterior, şi cu un nivel probabilistic semnificativ ales, se extrage t_α din tabelul 8.4.1. Mărimea sa se compară cu această valoare:

- dacă $D^2 < t_\alpha$, atunci ipoteza H_0 este acceptată;

- dacă $D^2 \geq t_\alpha$, atunci ipoteza H_0 este respinsă.

Pentru a mări acurateţea cu care legea probabilistică χ^2 aproximează comportamentul variabilei D^2, se recomandă ca, în limita posibilităţilor, submulţimile alese să fie echiprobabile, numărul lor să fie suficient de mare (cel puţin 3), dar sub 50, iar pentru fiecare submulţimi numărul de observaţii contorizate (cumulate) să depăşească valoarea 5 (la limită 3).

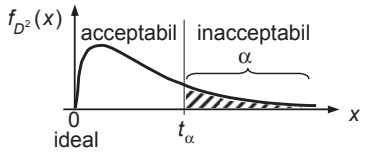


Figura 8.4.1: Determinarea deviaţiei critice, t_α

Tabelul 8.4.1: Deviaţii critice, t_α, funcţie de nivelul probabilistic semnificativ, α, şi de numărul de intervale, k

α	5%	1%	α	5%	1%
$k-1$	t_α		$k-1$	t_α	
1	3,84	6,63	12	21,03	26,22
2	5,99	9,21	13	22,36	27,69
3	7,81	11,35	14	23,69	29,14
4	9,49	13,28	15	25,00	30,58
5	11,07	15,09	16	26,30	32,00
6	12,59	16,81	17	27,59	33,41
7	14,07	18,48	18	28,87	34,81
8	15,51	20,09	19	30,14	36,19
9	16,92	21,67	20	31,41	37,57
10	18,31	23,21	25	37,65	44,31
11	19,68	24,76	30	43,77	50,89

Observaţii:

- Deviaţia critică poate fi obţinută şi printr-o aproximare Gauss a variabilei aleatoriii D^2 care reprezintă o sumă de variabile aleatorii (vezi teorema limită centrală).
- Atunci când un număr, r, de parametri ai distribuţiei presupuse sunt estimaţi pe baza rezultatelor observate (de exemplu: media, varianţa etc.), este recomandat ca variabila aleatorie D^2 să se considere o variabilă aleatorie χ^2 cu $k-r-1$ grade de libertate.

Aplicaţia 8.4.3

Luând în considerare informaţia din fişierul însoţitor `data2.txt`, care reprezintă momentele în care pachetele sunt depuse în interfaţa de ieşire a unui ruter, să se stabilească modelul matematic care descrie în mod adecvat procesul sosirilor, exprimat prin *număr de sosiri la fiecare 100 msec.*
Indicaţii:

1. Se transferă datele din fişier în vectorul `r`, folosind instrucţiunea `load`
2. Se fixează durata intervalului de referinţă, `lapse`, la valoarea de 100 msec
3. Se fixează durata observării fenomenului, `max_time` , prin majorarea celei mai mari înregistrări, înscrise în vectorul `r`, (vezi funcţia `max`) la un multiplu de `lapse` (vezi funcţia `ceil`)

4. Se defineşte vectorul `z`, al capetelor de intervale, care acoperă toată durata de observare
5. Se stabileşte numărul observat de sosiri, `arriv_no`, pe fiecare interval succesiv, în parte, folosind funcţia `histc`
6. Conţinutul vectorului `arriv_no` se reprezintă grafic cu funcţia `stem`, pentru a vizualiza realizarea particulară, observată, a procesului analizat
7. Se stabileşte plaja de variaţie, anume vectorul `k`, a numărului de sosiri pe un interval dat, cuprinsă între `0` şi valoarea maximă din vectorul `arriv_no`, incrementată cu 1
8. Se defineşte vectorul `k_interv_no`, ce reprezintă numărul **observat** de intervale, cu acelaşi număr de sosiri, şi se calculează elementele sale. De exemplu, conform figurii trasate anterior, în 4 intervale s-au înregistrat câte 2 sosiri.
9. Se calculează media, `m`, şi deviaţia standard, `s`, pe eşantionul de rezultate conţinut în vectorul `arriv_no` (vezi funcţiile `mean` şi `std`)
10. Luând în considerare drept ipoteze posibile distribuţiile teoretice Poisson, geometrică şi binomială, se calculează pentru fiecare valorile parametrilor asociaţi (vezi [Ioan]): `a_pois`, `p_geo`, `n_bin` şi `p_bin`
11. Se calculează vectorii probabilităţilor corespunzătoare acestor distribuţii teoretice, `ppois`, `pgeo` şi `pbin`, în punctele conţinute în vectorul `k` (vezi funcţia `pdf`).
12. Pentru fiecare distribuţie, se calculează vectorul numărului **aşteptat** de intervale cu acelaşi număr de sosiri, `narriv_pois`, `narriv_geo` şi `narriv_bin`
13. Pentru aprecierea calitativă a asemănării dintre rezultatele aşteptate şi cele observate se desenează, pe acelaşi grafic, histogramele corespunzătoare. Se foloseşte instrucţiunea `bar(k,p`, în care matricea `p` are drept coloane vectorii numerelor aşteptate şi observate de intervale cu acelaşi număr de sosiri
14. Pentru aprecierea cantitativă şi luarea deciziei finale se calculează statistica testului χ^2 pentru fiecare ipoteză în parte, `dpois`, `dgeo` şi `dbin`, conform relaţiei (8.4.2) şi se compară cu deviaţia critică (vezi tabelul 8.4.1)

8.5 Testul Kolmogorov-Smirnov

Spre deosebire de testul χ^2 care compară probabilităţile empirice ale unor intervale de genul (x_i, x_{i+1}), cu probabilitatea corespunzătoare, generate de funcţia densitate de probabilitate propusă, **testul Kolmogorov-Smirnov**, notat prin **K-S**, operează asupra diferenţelor ce apar între funcţia empirică, de distribuţie, obţinută pe baza unui eşantion de realizări şi funcţia analitică luată în considerare. Din acest motiv, testul K-S se pretează doar situaţiilor în care funcţia analitică de distribuţie este continuă.

Funcţia empirică de distribuţie a unei variabilei aleatorii, X, reprezintă, la rândul ei, o variabilă aleatorie şi este dată de relaţia:

$$\hat{F}_X(x) = \frac{\text{numărul rezultatelor ce sunt} \le x}{n} \qquad (8.5.1)$$

unde rezultatele aparţin unui eşantion de n observaţii independente, $x_1, x_2, \cdots, x_n$, ale variabilei aleatorii în cauză.

Figura 8.5.1 prezintă o funcţie empirică de distribuţie (linia în zigzag, ale cărui verticale, de fapt, nu există) şi mai multe funcţii analitice de distribuţie, propuse spre a caracteriza variabila aleatorie urmărită. După cum se poate observa, cele mai mici abateri apar când funcţia empirică, de distribuţie, este comparată cu curba centrală. Considerând suficientă comparaţia vizuală, putem concluziona că variabila aleatorie urmărită, X, şi variabila aleatorie Y_2 aparţin aceleiaşi familii de variabile aleatorii şi, deci, că funcţiile lor, analitice, de distribuţie sunt identice:

$$F_X(x) \equiv F_{Y_2}(x) \qquad (8.5.2)$$

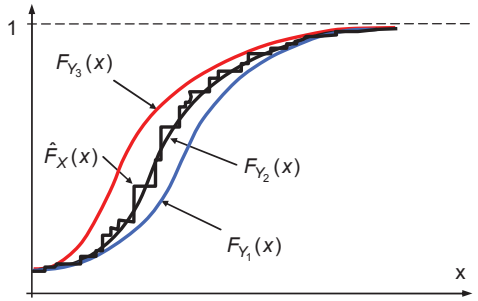


Figura 8.5.1: Funcţie empirică de distribuţie şi compararea ei cu trei funcţii analitice de distribuţie - exemplu

Deseori, datorită subiectivităţii sale, compararea vizuală a curbelor ne poate înşela. Testul K-S vine să elimine acest neajuns, diferenţele dintre $\hat{F}_X(x)$ şi $F_Y(x)$ fiind evaluate cantitativ, pe baza următoarelor **statistici** (variabile aleatorii la rândul lor):

$$D_n^+ = \sqrt{n}\left\{\max_{-\infty<x<+\infty}\left(\hat{F}_X(x) - F_Y(x)\right)\right\}$$
$$D_n^- = \sqrt{n}\left\{\max_{-\infty<x<+\infty}\left(F_Y(x) - \hat{F}_X(x)\right)\right\} \qquad (8.5.3)$$

Mărimea D_n^+ reprezintă cea mai mare deviaţie (vezi figura 8.5.2) când $\hat{F}_X(x) > F_Y(x)$, iar valoarea D_n^- corespunde celei mai mari deviaţii când

$\hat{F}_X(x) < F_Y(x)$. Apariţia factorului $\sqrt{n}$ este motivată de faptul că deviaţia standard a variabilei aleatorii $\hat{F}_X(x)$ faţă de $F_Y(x)$ este proporţională cu $1/\sqrt{n}$ şi deci că, prin această operaţie, variabilele aleatorii D_n^+ şi D_n^- devin independente faţă de n.

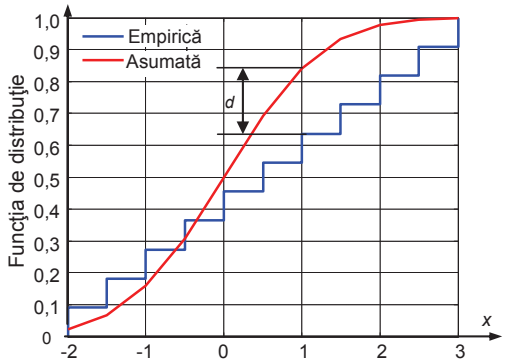


Figura 8.5.2: Exemplu de deviaţie între funcţiile de distribuţie empirică şi ipotetică, folosite în testul K-S

Testul Kolmogorov-Smirnov poate fi:

- **de o singură parte** (***one-sided test***), când se are în vedere doar una din cele două statistici, D_n^+ sau D_n^-,
- **de ambele părţi** (***two-sided test***), când se ia în considerare statistica:

$$D_n = \sqrt{n} \cdot \left| \max_{-\infty<x<+\infty}\left(\hat{F}_X(x) - F_Y(x)\right)\right| = \max\left(D_n^+, D_n^-\right) \qquad (8.5.4)$$

Luând în considerare o valoare particulară, d şi o ipoteză, H_0, privind distribuţia asumată, se poate calcula o valoare, numită ***"valoare p"*** (***p value***), care exprimă gradul de credibilitate a ipotezei. Concret, *valoarea* p este probabilitatea ca, în condiţiile considerării ipotezei H_0 adevărate, statistica unui eşantion oarecare de rezultate, D, să fie egală sau să depăşească realizarea particulară, d, a eşantionului curent de rezultate, adică:

$$\textit{valoarea } p = P[D \ge d] \qquad (8.5.5)$$

Putem considera că valorile mari ale *"valorii p"* presupun că abaterea rezultată se datorează aleatoriului, iar valorile mici au drept cauză considerarea unei ipoteze H_0 falsă. Prin urmare, având în vedere un ***nivel probabilistic semnificativ***, α, şi ţinând cont de tipul de test K-S desfăşurat:

- ipoteza H_0 se respinge dacă:
  - *valoarea* $p \le \alpha$, pentru teste de o singură parte,
  - *valoarea* $p \le \alpha/2$, pentru teste de ambele părţi,
- ipoteza H_0 se acceptă în caz contrar (vezi figura 8.8.3).

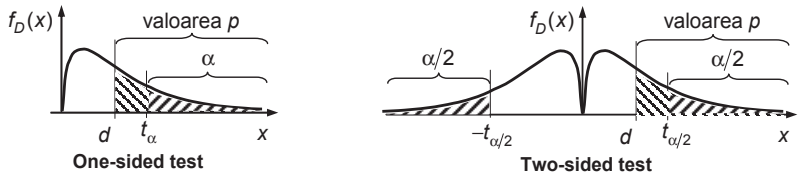


Figura 8.5.3: Cazuri de acceptare a ipotezei H_0

Mărimea t_α reprezintă **deviaţia critică** a cărei valoare depinde de nivelul probabilistic semnificativ şi de dimensiunea eşantionului de rezultate, conform tabelului 8.5.1. Luând în considerare această mărime:

- ipoteza H_0 se respinge dacă $d \ge t_\alpha$
- ipoteza H_0 se acceptă în caz contrar.

Decizia, dacă acceptăm sau nu ipoteza că variabila aleatorie urmărită, X, se poate identifica cu variabilă aleatorie propusă, Y, se ia comparând mărimile calculate, D_n^+ şi D_n^- cu valorile ***deviaţiilor critice*** oferite în tabelul 8.5.1, pentru diferite **nivele** probabilistice **semnificative**, α, şi diverse dimensiuni, n, ale eşantionului.

Din punct de vedere practic, relaţiile (8.5.3) şi (8.5.4) sunt inutilizabile, obligându-ne să căutăm extreme pe un domeniu infinit ($-\infty < x < +\infty$). Ţinând cont, însă, că $F_Y(x)$ este monoton crescătoare şi că $\hat{F}_X(x)$ este crescătoare în trepte, putem folosi următorul procedeu de evaluare a mărimilor D_n^+ şi D_n^-:

- Se obţin rezultatele $x_1, x_2, \cdots, x_n$;
- Rezultatele sunt ordonate crescător, obţinându-se secvenţa: $x'_1, x'_2, \cdots, x'_n$;
- Mărimile D_n^+ şi D_n^- se obţin aplicând formulele:

$$D_n^+ = \sqrt{n} \cdot \left[\max_{1 \le j \le n} \left(\frac{j}{n} - F_Y(x'_j) \right) \right]$$
$$D_n^- = \sqrt{n} \cdot \left[\max_{1 \le j \le n} \left(F_Y(x'_j) - \frac{j-1}{n} \right) \right] \qquad (8.5.6)$$

Pentru ca testul K-S să fie edificator trebuie ca numărul de observaţii, n, să se aleagă suficient de mare, uzual 1000, astfel încât funcţiile empirice să fie

cât mai "fine", cu cât mai multe trepte, ceea ce asigură o cantitate suficientă de valori ale diferenţelor, ce pot fi comparate.

Tabelul 8.5.1: Deviaţii critice, t_α, în cazul testului Kolmogorov-Smirnov (*one-sided test*)

n	0.1	0.05	0.025	0.01	0.005
1	0.9000	0.9500	0.9750	0.9900	0.9950
2	0.6838	0.7764	0.8419	0.9000	0.9293
3	0.5648	0.6360	0.7076	0.7846	0.8290
4	0.4927	0.5652	0.6239	0.6889	0.7342
5	0.4470	0.5094	0.5633	0.6272	0.6685
6	0.4104	0.4680	0.5193	0.5774	0.6166
7	0.3815	0.4361	0.4834	0.5384	0.5758
8	0.3583	0.4096	0.4543	0.5065	0.5418
9	0.3391	0.3875	0.4300	0.4796	0.5133
10	0.3226	0.3687	0.4092	0.4566	0.4889
11	0.3083	0.3524	0.3912	0.4367	0.4677
12	0.2958	0.3382	0.3754	0.4192	0.4490
13	0.2847	0.3255	0.3614	0.4036	0.4325
14	0.2748	0.3142	0.3489	0.3897	0.4176
15	0.2659	0.3040	0.3376	0.3771	0.4042
16	0.2578	0.2947	0.3273	0.3657	0.3920
17	0.2504	0.2863	0.3180	0.3553	0.3809
18	0.2436	0.2785	0.3094	0.3457	0.3706
19	0.2373	0.2714	0.3014	0.3369	0.3612
20	0.2316	0.2647	0.2941	0.3287	0.3524
21	0.2262	0.2586	0.2872	0.3210	0.3443
22	0.2212	0.2528	0.2809	0.3139	0.3367
23	0.2165	0.2475	0.2749	0.3073	0.3295
24	0.2120	0.2424	0.2693	0.3010	0.3229
25	0.2079	0.2377	0.2640	0.2952	0.3166
26	0.2040	0.2332	0.2591	0.2896	0.3106
27	0.2003	0.2290	0.2544	0.2844	0.3050
28	0.1968	0.2250	0.2499	0.2794	0.2997
29	0.1935	0.2212	0.2457	0.2747	0.2947
30	0.1903	0.2176	0.2417	0.2702	0.2899
31	0.1873	0.2141	0.2379	0.2660	0.2853
32	0.1844	0.2108	0.2342	0.2619	0.2809
33	0.1817	0.2077	0.2308	0.2580	0.2768
34	0.1791	0.2047	0.2274	0.2543	0.2728
35	0.1766	0.2018	0.2242	0.2507	0.2690
36	0.1742	0.1991	0.2212	0.2473	0.2653
37	0.1719	0.1965	0.2183	0.2440	0.2618
38	0.1697	0.1939	0.2154	0.2409	0.2584
39	0.1675	0.1915	0.2127	0.2379	0.2552
40	0.1655	0.1891	0.2101	0.2349	0.2521
> 40	$1.07/\sqrt{n}$	$1.22/\sqrt{n}$	$1.36/\sqrt{n}$	$1.52/\sqrt{n}$	$1.63/\sqrt{n}$

Aplicaţia 8.5.1

Observarea unei variabile aleatorii X a permis obţinerea următorului eşantion de rezultate:

0,869; 0,189; 0,678; 0,259; 0,718; 0,098; 0,318; 0,733; 0,442; 0,141; 0,236; 0,017; 0,434; 0,162; 0,414; 0,354; 0,772; 0,442; 0,302; 0,693.

Considerând ca ipoteză: "*variabila aleatorie X urmează o lege de distribuţie uniformă în intervalul* $[0,1]$" :

i) *Să se traseze*, pe acelaşi sistem de axe, graficele pentru funcţia empirică de distribuţie şi pentru funcţia analitică de distribuţie propusă;

ii) *Să se verifice*, prin testul K-S, dacă ipoteza avută în vedere poate fi acceptată la un nivel semnificativ de 95%.

* * *

Aplicaţia 8.5.2

Scrieţi un program Matlab care să permită compararea grafică a funcţiilor de distribuţie, ***cdf***, şi densitate a probabilităţilor, ***pdf***, în cazul distribuţiilor gaussiene, gama şi exponenţiale, folosind acelaşi valori pentru parametrii specifici acestor distribuţii. Compararea se va face ilustrând în aceeaşi figură curbele cdf, respectiv pdf, ce corespund celor 3 distribuţii analizate. Analizaţi influenţa mărimilor alese pentru parametri asupra valorilor probabilităţilor din fiecare distribuţie.

Indicaţii:

1. se defineşte spaţiul realizărilor: `x = ...;`
2. se calculează valorile funcţiilor de distribuţie şi de densitate de probabilitate pentru variabile de tipul Gauss, gama şi exponenţial având aceeaşi medie, $m = 4$, şi următorii coeficienţi de variaţie: $1,\ 2,\ 4$ (vezi `help cdf` şi `help pdf`).

se trasează pe acelaşi grafic (vezi `help plot` şi `help hold`) curbele privind funcţiile de distribuţie, respectiv de densitate de probabilitate, considerate.

* * *

Aplicaţia 8.5.3

Considerând experimentul *"aruncarea cu moneda de 20 de ori"* să se verifice ipoteza *"feţele monedei sunt echiprobabile"* dacă:

- nivelul probabilistic semnificativ considerat este $\alpha = 0,05$,
- s-a obţinut de 14 ori aceeaşi faţă,
- statistica avută în vedere este $D = \max(H\_\text{no}, T\_\text{no}) - 10$, unde H_no este numărul de feţe *"Head"*, iar T_no este numărul de feţe *"Tail"*.

Reluaţi aplicaţia dacă aceeaşi faţă apare de 15 ori.

Indicaţie:

- Se calculează *"valoarea p"*, notată în program `p_val`, ţinând cont că o aruncare individuală cu banul este caracterizată, din punct de vedere probabilistic, de o distribuţie Bernoulli de parametru p, care reprezintă şansa apariţiei unei anumite feţe, iar aruncarea repetată cu banul are ca model probabilistic distribuţia binomială, de parametri: n numărul de repetări şi p cu aceeaşi semnificaţie ca mai sus.
- Se compară `p_val` obţinută cu nivelul probabilistic semnificativ considerat şi se ia decizia corespunzătoare privind credibilitatea ipotezei.

* * *

Aplicaţia 8.5.4

Stabiliţi funcţia de distribuţie a probabilităţilor pentru durata de ocupare a unui canalul de transmisiuni. Rezultatele observaţilor efectuate asupra canalului sunt conţinute în fişierul însoţitor `data1.txt`. Realizaţi analiza datelor experimentale cu ajutorul testului Kolmogorov-Smirnov, considerând ca funcţii etalon distribuţiile Gauss-normală, exponenţială şi gama. Reprezentaţi grafic curbele corespunzătoare distribuţiei empirice şi distribuţiilor etalon pentru o apreciere calitativă, prin vizualizarea "potrivirilor". Pentru decizia cantitativă şi decisivă se va utiliza funcţia Matlab `kstest` pentru diferite niveluri semnificative de decizie. Comentaţi rezultatele obţinute.

Indicaţii:

1. Se încărcă datele experimentale, folosind instrucţiunea `r = load ('data1.txt')`, unde `r` reprezintă vectorul valorilor înscrise în fişierul `data1.txt` (vezi lucrarea IT_1, de laborator, din rubrica "Ingineria Traficului" de la adresa http://www.comm.pub.ro/users/~lucian/ro/);

2. Se desenează distribuţia empirică corespunzătoare vectorului `r` (vezi `cdfplot, grid`);
3. Se calculează valorile caracteristice ale eşantionului de valori cu ajutorul următoarelor funcţii: `mean` şi `std`. Aceste valori se atribuie parametrilor distribuţiilor etalon considerate,
4. Pentru ca testul K-S să fie corect realizat, compararea valorilor experimentale cu cele teoretice trebuie să se facă pentru acelaşi valori cuprinse în vectorul `r`, care suplimentar trebuie (vezi `help kstest`) să fie ordonate în sens crescător (vezi `help sort`). Se obţine astfel vectorul `x`;
5. Se construiesc vectorii care conţin valorile funcţiilor teoretice de distribuţie (CDF) în punctele cuprinse în `x` (vezi `cdf`);
6. Se trasează suprapus peste curba anterioară cele trei funcţii teoretice de distribuţie pentru analiza comparativă;
7. Se execută testul K-S, pentru fiecare CDF etalon, considerând implicit nivelul semnificativ ($\alpha = 0,05$), folosind instrucţiunea :

```
[h,p,d,ta] = kstest(x,[x y])
```

în care:

`x` reprezintă vectorul valorilor ordonate ale eşantionului, iar
`y` conţine valorile funcţiei CDF în punctele conţinute în vectorul `x`.

capitolul 9

Procese aleatorii

9.1 Noţiuni introductive

În anumite experimente aleatorii, realizările sunt funcţie de timp, spaţiu sau alte coordonate. Astfel, de exemplu, într-un sistem de servire cu aşteptare numărul clienţilor variază în timp, iar intensitatea unei imagini video se modifică de la un punct (*pixel*) la altul. Asemenea experimente se specifică, de obicei [Pap91], prin realizările lor, $\xi \in \mathrm{S}$, prin ansamblul de evenimente definite în S şi prin probabilităţile acestor evenimente.

Deoarece realizările ξ ale unui experiment se desfăşoară într-un anumit spaţiu, notat de exemplu cu I (***mulţimea de indexare***), ele pot fi descrise analitic prin intermediul unei familii de funcţii, identificată prin intermediul notaţiei $\{X(t,\xi);\ t \in \mathrm{I}\}$.

Figura 9.1 prezintă trei realizări particulare ale unui experiment aleatoriu cu desfăşurare în timp (în acest caz, mulţimea I este axa timpului). Aşa cum se poate observa, pentru oricare moment t_k, fixat, $X(t_k,\xi)$ este o variabilă aleatorie, motiv pentru care se poate considera că $X(t,\xi)$ este o înlănţuire de variabile aleatorii indexate, în acest caz, după t.

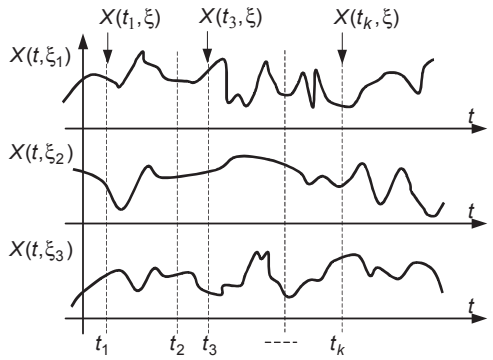


Figura 9.1: Realizări de experimente aleatorii cu desfăşurare în timp

O astfel de înlănţuire poartă denumirea de ***proces aleatoriu***, iar graficul funcţie $X(t,\xi)$ pentru o anumită realizare se numeşte ***realizare*** (***traiectorie***) a procesului aleatoriu. Dacă asupra unui proces aleatoriu se pot face previziuni statistice, atunci el este numit **proces *stocastic***. În continuare, pentru simplificarea scrierii, se vor folosi notaţiile:

$$X(t,\xi)=X(t) \text{ şi } X(t_k,\xi)=X(t_k)$$

Aplicaţia 9.1.1

Fie un proces aleatoriu, $X(t,\xi)$, care constă în recepţionarea din linia de comunicaţie a patru simboluri binare (0 sau 1), consecutive.

a) Ştiind că debitul transmisiei este de 100 kbiţi/sec, să se reprezinte grafic, specificând pe axa timpului momentele care separă simbolurile, trei realizări posibile obţinute în urma repetării procesului considerat (figura 9.1). Ca origine a axelor se va considera momentul începerii recepţionării primului simbol.

b) Câte elemente are spaţiul realizărilor procesului aleatoriu studiat?

c) Care este spaţiul realizărilor variabilei aleatorii $X(\theta)$ asociate procesului analizat, unde momentul θ este acelaşi faţă de începutul primului simbol pentru toate realizările posibile?

* * *

La acest nivel general de descriere a proceselor aleatorii, singurele elemente pe baza cărora se poate face o clasificare sunt caracteristicile domeniului în care variabilele aleatorii indexate, $X(t_k)$, $\forall t_k$, iau valori, precum şi ale mulţimii I pe parcursul căruia se desfăşoară acesta. În raport cu aceste considerente:

• un ***proces*** este ***discret***, dacă mulţimea I este finit sau infinit numărabilă, sau este ***continuu***, în caz contrar;

• un ***proces*** este ***evaluat discret***, dacă spaţiul realizărilor variabilelor aleatorii indexate este finit sau infinit numărabil, sau este ***evaluat continuu***, în caz contrar.

Modul de combinare al celor două însuşiri este prezentat în figura 9.2.

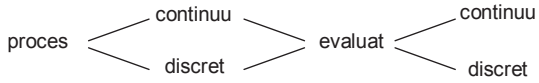


Figura 9.2: Clasificarea generală a proceselor aleatorii

În general, noţiunea de proces evaluat discret se substituie cu cea de ***lanţ*** (sau ***serie***). În acest caz, vorbim de ***lanţuri continue*** sau ***discrete*** ale căror valori poartă denumirea de ***stări*** ale procesului. De asemenea, este util să evidenţiem prin notaţie natura discretă a spaţiului de desfăşurare a procesului, mulţimea I, lanţurile sau procesele în cauză notându-se cu X_n.

Aplicaţia 9.1.2

Stabiliţi din ce categorie fac parte următoarele procese:

a) cel prezentat în figura 9.1;
b) cel analizat în aplicaţia 9.1.1;
c) sondarea periodică, de exemplu din 5 în 5 minute a numărului de conexiuni stabilite în interiorul unui sistem de comutaţie;
d) eşantionarea unui semnal telefonic.

9.2 Specificarea unui proces aleatoriu

În general, [Mih78], un proces aleatoriu este descris complet prin intermediul unui ansamblu (colecţii) de funcţii de distribuţie de legătură, corespunzătoare tuturor combinaţiilor posibile ale variabilelor indexate, aferente, de forma:

$$F_{X(t_1),X(t_2),\cdots,X(t_k)}(x_1,x_2,\cdots,x_k)=P\left[X(t_1)\le x_1,\cdots,X(t_k)\le x_k\right],\forall k \quad (9.2.1)$$

Aceste colecţii se împart în distribuţii de ordinul 1 pentru $k=1$, distribuţii de ordinul 2 pentru $k=2$ etc.

Dacă un proces este evaluat discret, el poate fi specificat şi printr-o colecţie de probabilităţi elementare de legătură, acestea fiind de forma:

$$p_{X(t_1),X(t_2),\cdots,X(t_k)}(x_1,x_2,\cdots,x_k)=P\left[X(t_1)=x_2\ \&\cdots\&\ X(t_k)=x_k\right],\forall k \quad (9.2.2)$$

Dacă este evaluat continuu, procesul poate fi specificat şi printr-o colecţie de funcţii de densitate de probabilitate de legătură, notate sub forma:

$$f_{X(t_1),X(t_2),\cdots,X(t_k)}(x_1,x_2,\cdots,x_k),\ \forall k \quad (9.2.3)$$

O specificare parţială a proceselor se poate realiza cu ajutorul *valorilor matematice caracteristice* (***momente***). Dintre acestea, cele mai folosite sunt:

• ***media*** (speranţa matematică sau momentul de ordinul 1):

$$\mathrm{E}\left[X(t)\right]=\int_{-\infty}^{+\infty}x\cdot f_{X(t)}(x)\mathrm{d}x \overset{\text{notat}}{=} m_X(t) \quad (9.2.4)$$

• ***varianţa*** (momentul centrat de ordinul 2):

$$\mathrm{VAR}\left[X(t)\right]=\mathrm{E}\left[\left(X(t)-m_x(t)\right)^2\right]\overset{\text{notat}}{=}\sigma_X^2(t) \quad (9.2.5)$$

• ***autocorelaţia*** :

$$R_X(t_1,t_2)=\mathrm{E}\left[X(t_1),X(t_2)\right]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\cdot f_{X(t_1),X(t_2)}(x,y)\,\mathrm{d}x\,\mathrm{d}y \quad (9.2.6)$$

• ***autocovarianţa*** :

$$\begin{aligned}C_X(t_1,t_2)&=\mathrm{E}\left[\left\{X(t_1)-m_X(t_1)\right\}\cdot\left\{X(t_2)-m_X(t_2)\right\}\right]\\&=R_X(t_1,t_2)-m_X(t_1)\cdot m_X(t_2)\end{aligned} \quad (9.2.7)$$

• ***coeficientul de autocorelaţie*** :

$$\rho_X(t_1,t_2)=\frac{C_X(t_1,t_2)}{\mathrm{STD}\left[X(t_1)\right]\cdot\mathrm{STD}\left[X(t_2)\right]} \quad (9.2.8)$$

Aplicaţia 9.2.1

Stabiliţi numărul membrilor care alcătuiesc colecţia caracteristică de distribuţii a unui proces discret, ce se desfăşoară pe o mulţime de indexare finită.

* * *

Exemplul cel mai simplu **de proces aleatoriu** este o familie, X_n, de variabile aleatorii independente şi identic distribuite (au aceeaşi funcţie de distribuţie, deci aceeaşi medie, m, aceeaşi varianţă, σ^2 etc.) reprezintă un ***proces aleatoriu discret independent identic distribuit*** (**i.i.d.**). În acest caz, funcţiile de distribuţie de legătură se pot exprima sub formă de produs:

$$F_{X_{n_1},X_{n_2},\cdots,X_{n_k}}(x_1,x_2,\cdots,x_k)=P\left[X_{n_1}\le x_1\right]\cdot P\left[X_{n_2}\le x_2\right]\cdot\ldots\cdot P\left[X_{n_k}\le x_k\right] \quad (9.2.9)$$

unde: $P\left[X_{n_i}\le x_i\right]=F_{X_i}(x_i)$ este funcţia de distribuţie individuală, $i=\overline{1,k}$.

Proprietăţile importante ale unui astfel de proces sunt:

- **media**: $m_X(n)=\mathrm{E}\left[X_n\right]=m$, $\forall n$, unde m este o constantă

- **varianţa**: $\mathrm{VAR}\left[X_n\right]=\sigma^2$

- **autocorelaţia**: $R_X\left[n_1,n_2\right]=\begin{cases}m^2 & n_1\ne n_2\\ m^2+\sigma^2 & n_1=n_2\end{cases}$

- **autocovarianţa**:

$$C_X(n_1,n_2)=\mathrm{E}\left[(X_{n_1}-m)\cdot(X_{n_2}-m)\right]=\begin{cases}0 & \text{dacă } n_1\neq n_2\\ \sigma^2 & \text{dacă } n_1=n_2\end{cases}$$

- **coeficientul de corelaţie**: $\rho_X(n_1,n_2)=\begin{cases}0 & \text{pentru } n_1\neq n_2\\ 1 & \text{pentru } n_1=n_2\end{cases}$

Aplicaţia 9.2.2 **(procesul Bernoulli)**

Fie X_n o secvenţă de variabile aleatorii i.i.d. după o *lege Bernoulli* de parametru $p=0{,}4$. Reprezentaţi grafic o realizare posibilă a procesului pe parcursul a 5 valori consecutive ale spaţiului de desfăşurare. Determinaţi care este probabilitatea ca într-o realizare a acestui proces să se obţină combinaţia de cifre consecutive:

a) 1001; b) $1,\#,\#,\#,0$, unde simbolul # semnifică orice valoare posibilă.

9.3 Staţionaritatea proceselor aleatorii

Un proces aleatoriu $X(t)$ este prin definiţie (*strict*) ***staţionar***, dacă distribuţia de legătură a oricărui set de variabile aleatorii indexate, $X(t_i)$, unde $i=\overline{1,k}$, $(\forall k)$ depinde (eventual) doar de distanţa relativă între momentele sondării şi nu de poziţia lor (absolută) în domeniul de indexare, adică:

$$F_{X(t_1),X(t_2),\cdots,X(t_k)}(x_1,x_2,\cdots,x_k)=F_{X(t_1+\tau),\cdots,X(t_k+\tau)}(x_1,x_2,\cdots,x_k) \quad (9.3.1)$$

Cazurile particulare importante, derivate din relaţia de definiţie (9.3.1) precizează că:

- distribuţia de ordinul 1 a unui proces aleatoriu staţionar este independentă de timp (este aceeaşi indiferent unde în domeniul de indexare), adică:

$$F_{X(t)}(x)=F_{X(t+\tau)}(x) \text{ pentru } \forall\, t,\tau \quad (9.3.2)$$

- distribuţie de legătură de ordinul 2 a unui proces staţionar depinde doar de diferenţa de timp între momentele de observare, adică:

$$F_{X(t_1)X(t_2)}(x_1,x_2)=F_{X(0)X(t_2-t_1)}(x_1,x_2) \text{ pentru } \forall\, t_1,t_2 \quad (9.3.3)$$

Relaţiile (9.3.2) şi (9.3.3) au drept consecinţe câteva proprietăţi utile în a verifica dacă un anumit proces supus analizei ***nu este (strict) staţionar***. Aceste proprietăţi (prin urmare *necesare*, *dar nu şi suficiente* pentru a caracteriza un proces strict staţionar) sunt:

➢ **media** şi **varianţa** unui proces aleatoriu staţionar sunt constante, independente de timp, anume:

$$m_x(t)=m \text{, pentru } \forall t \quad (9.3.4)$$

$$\mathrm{VAR}\left[X(t)\right]=\mathrm{E}\left[\left(X(t)-m\right)^2\right]=\sigma^2 \text{, pentru } \forall t \quad (9.3.5)$$

➢ **autocorelaţia** şi **autocovarianţa** şi, implicit, **coeficientul de corelaţie** al unui proces aleatoriu staţionar depind (eventual) doar de diferenţa relativă dintre momentele t_1 şi t_2 şi nu de valorile particulare ale acestora, deci pentru $\forall\, t_1,t_2$:

$$R_X(t_1,t_2)=R_X(t_2-t_1) \quad (9.3.6)$$

$$C_X(t_1,t_2)=C_X(t_2-t_1) \quad (9.3.7)$$

$$\rho_X(t_1,t_2)=\rho_X(t_2-t_1) \quad (9.3.8)$$

Aplicaţia 9.3.1

Fie o secvenţă, X_n, de variabile aleatorii i.i.d. după o lege Bernoulli de parametru p (după o lege exponenţială de parametru λ ş.a.m.d.).

a) arătaţi că un astfel de proces este staţionar;

b) calculaţi mărimile caracteristice: media, varianţa, autocorelaţia, autocovarianţa şi coeficientul de corelaţie;

c) calculaţi distribuţia condiţionată: $P\left[X_{n_2}\le x_2 \,\middle|\, X_{n_1}=x_1\right]$, unde x_1 şi x_2 sunt două realizări posibile ale variabilelor considerate. Interpretaţi rezultatele.

Rezolvare:

a) Indiferent de lege, relaţia de definiţie (9.3.1) este verificată deoarece variabilele aleatorii corespunzătoarea diverselor momente de sondare a procesului sunt *independente* şi *identic distribuite* conform enunţului.

b) În cazul legii Bernoulli se obţine: $\square\ m_X(n)=p$; $\square\ \mathrm{VAR}\left[X_n\right]=p\cdot(1-p)$

$$\square\ R_X(n_1,n_2)=\mathrm{E}\left[X_{n_1}\cdot X_{n_2}\right]\overset{\text{independenţa variabilelor}}{=}\mathrm{E}\left[X_{n_1}\right]\cdot\mathrm{E}\left[X_{n_2}\right]\overset{\text{identitatea variabilelor}}{=}\begin{cases}p^2 & \text{pentru } n_1\neq n_2\\ p & \text{pentru } n_1=n_2\end{cases}$$

$$\square\ C_X(n_1,n_2)=R_X(n_1,n_2)-m_X(n_1)\cdot m_X(n_2)=\begin{cases}0 & \text{pentru } n_1\neq n_2\\ p\cdot(1-p) & \text{pentru } n_1=n_2\end{cases}$$

$$\square\ \rho_X(n_1,n_2)=\begin{cases}0 & \text{pentru } n_1\neq n_2\\ 1 & \text{pentru } n_1=n_2\end{cases}$$

c) Pentru probabilitatea condiţionată cerută se obţine:

$$P\left[X_{n_2}\le x_2 \,\middle|\, X_{n_1}=x_1\right]=\frac{P\left[\left\{X_{n_2}\le x_2\right\}\&\left\{X_{n_1}=x_1\right\}\right]}{P\left[X_{n_1}=x_1\right]}$$

$$\overset{\text{independenţa variabilelor}}{=}\frac{P\left[X_{n_2}\le x_2\right]\cdot P\left[X_{n_1}=x_1\right]}{P\left[X_{n_1}=x_1\right]}=P\left[X_{n_2}\le x_2\right]$$

Interpretare: Secvenţa analizată este necorelată ($C_X(n_1,n_2)=0$), ceea ce face ca şi distribuţia într-un anumit punct de observare, n_2, să nu depindă de ceea ce s-a observat într-un punct anterior, n_1.

Aplicaţia 9.3.2

Verificaţi dacă un proces aleatoriu: a) discret, b) continuu, de tip sumă este staţionar.

Rezolvare: Contrazicerea unei afirmaţii este, în general, un demers mult mai facil decât confirmarea sa, fiind suficientă identificarea unei singure proprietăţi derivate din enunţ care nu se respectă. Astfel, în cazul procesului aleatoriu discret de tip sumă, media corespunzătoarea are expresia:

$$m_S(n)=\mathrm{E}\left[S_n\right]=\mathrm{E}\left[X_1+X_2+\ldots+X_n\right]\overset{\text{i.i.d.}}{=}n\cdot m$$

care, prin faptul că este crescătoare cu n, conduce la concluzia că procesul nu este staţionar.

* * *

Multe din procesele aleatorii întâlnite în diverse aplicaţii practice nu se încadrează în categoria proceselor staţionare. În aceste condiţii este util să se verifice dacă ele aparţin unei clase mai largi, precum cea a *proceselor asimptotic staţionare* sau chiar cea a *proceselor staţionare în sens larg*, pentru care există, încă, metode analitice de analiză şi prelucrare (de exemplu, tehnicile de prelucrare a semnalelor sau utilizarea formulelor bazate pe valori medii):

- Un ***proces aleatoriu*** este ***asimptotic staţionar*** dacă:

$$F_{X(t_1),X(t_2),\cdots,X(t_k)}(x_1,x_2,\cdots,x_k)\overset{t_1\to\infty}{=}F_{X(t_1+\tau),X(t_2+\tau),\cdots,X(t_k+\tau)}(x_1,x_2,\cdots,x_k)\quad(9.3.8)$$

pentru orice secvenţă $t_1<t_2<\cdots<t_k$ în număr oarecare de termeni ($\forall k$).

- Un ***proces aleatoriu*** este ***staţionar în sens larg*** dacă respectă următoarele relaţii:

$$\begin{aligned}&m_x(t)=m,\forall t\\&R_X(t_1,t_2)=R_X(t_2-t_1),\forall t_1,t_2\\&C_X(t_1,t_2)=C_X(t_2-t_1).\forall t_1,t_2\end{aligned}\qquad(9.3.9)$$

Aplicaţia 9.3.3

Fie procesul aleatoriu discret X_n compus prin întreţeserea a două secvenţe independente de variabile aleatorii.

Prima secvenţă *ocupă* poziţiile pare ale procesului ($n=2k$) fiind alcătuită din variabile aleatorii i.i.d. cu spaţiul realizărilor $\{-1,1\}$ şi probabilităţi elementare $p_{-1}=p_1=0{,}5$, iar a doua secvenţă ocupă poziţiile impare ale procesului ($n=2k+1$) fiind reprezentată de variabile aleatorii i.i.d. cu spaţiul realizărilor

$\{-3,1/3\}$ şi probabilităţi elementare $p_{-3}=1/10$ şi $p_{1/3}=9/10$.

a) Reprezentaţi grafic o realizare posibilă a procesului luând în considerare cel puţin patru momente (poziţii) succesive.

b) Stabiliţi genul de staţionaritate care caracterizează procesul analizat.

Rezolvare b): Deoarece probabilităţile elementare de ordinul 1 diferă de la n par la n impar, $p_{X_{2k}}(x)\neq p_{X_{2k+1}}(x)$, înseamnă că procesul nu este strict staţionar. Cu toate acestea: $m_X(n)=\ldots=0$, $\forall n$ şi

$$C_X(i,j)=R_X(i,j)=\begin{cases}\mathrm{E}\left[X_i\right]\cdot\mathrm{E}\left[X_j\right]=0 & \text{pentru } i\neq j\\ \mathrm{E}\left[X_i^{\,2}\right]=1 & \text{pentru } i=j\end{cases}$$

ceea ce arată că procesul este staţionar în sens larg.

* * *

Informaţii suplimentare referitoare la procesele staţionare, cel puţin în sens larg, se pot obţine prin intermediul autocorelaţiei, $R_X(t_2-t_1)$, notată şi prin $R_X(\tau)$, cu $\tau=t_2-t_1$. Astfel, următoarea relaţie, obţinută prin aplicarea inegalităţii Markov:

$$\begin{aligned}P\left[\left|X(t+\tau)-X(t)\right|>\varepsilon\right]&=P\left[\left(X(t+\tau)-X(t)\right)^2>\varepsilon^2\right]\\&\leq\frac{\mathrm{E}\left[\left(X(t+\tau)-X(t)\right)^2\right]}{\varepsilon^2}=\frac{2\left(R_X(0)-R_X(\tau)\right)}{\varepsilon^2}\end{aligned}\qquad(9.3.10)$$

reprezintă un mijloc de măsurare a "vitezei" cu care un proces aleatoriu se schimbă între două momente diferite, t şi $t+\tau$.

De asemenea, ***puterea medie*** (*momentul de ordinul doi*) a unui proces aleatoriu este dată de relaţia:

$$\mathrm{E}\left[X^2(t)\right]=R_X(0)=\int_{-\infty}^{+\infty}S_X(f)\,\mathrm{d}f\qquad(9.3.11)$$

unde $S_x(f)$ reprezintă "densitatea de putere" a procesului $X(t)$. Cea de-a doua exprimare rezultă ţinând cont că $S_x(f)$ este transformata Fourier a autocorelaţiei [MDS01], fiind astfel valabile egalităţile:

$$\begin{aligned}S_X(f)&=\mathrm{F}\left\{R_X(\tau)\right\}=\int_{-\infty}^{+\infty}R_X(\tau)\exp(-j2\pi f\tau)\,\mathrm{d}\tau\\R_X(\tau)&=\mathrm{F}^{-1}\left\{S_X(f)\right\}=\int_{-\infty}^{+\infty}S_X(f)\exp(-j2\pi f\tau)\,\mathrm{d}f\end{aligned}\qquad(9.3.12)$$

Aplicaţia 9.3.4

Fie un sistem discret, cauzal, al cărui semnal de ieşire este dat de următoarea ecuaţie cu diferenţe finite: $Y_n=\alpha\cdot Y_{n-1}+X_n$, $|\alpha|<1$. Considerând că

X_n este o secvenţă de variabile aleatorii i.i.d., de medie 0 şi de varianţă σ_x^2, arătaţi că:

a) Sistemul este liniar.

b) Sistemul este invariant în timp.

c) Densitatea de putere a procesului Y_n are expresia:

$$S_Y(f) = \sigma_x^2 \Big/ \left(1 + \alpha^2 - 2\alpha \cdot \cos 2\pi f\right)$$

Indicaţie:

a) Răspunsul $y(t)$ al unui sistem la un semnal de intrare $x(t)$ este obţinut, la modul general, în urma unei transformări exprimate analitic prin $y(t) = T[x(t)]$.

Un sistem este liniar dacă prezintă proprietatea de superpoziţie, adică:

$$T[a \cdot x_1(t) + b \cdot x_2(t)] = a \cdot T[x_1(t)] + b \cdot T[x_2(t)]$$

b) Un sistem este invariant în timp dacă răspunsul la $x(t-\tau)$ este $y(t-\tau)$.

c) Un sistem este cauzal dacă răspunsul la momentul t depinde doar de valorile semnalului de intrare, generate până în acel moment, adică, în cazul discret, dacă $h_n = 0$, pentru $n < 0$, unde $h_n = T[\delta_n]$ este **funcţia pondere** a sistemului, ca răspuns la semnalul impuls: $\delta_n = \begin{cases} 1 & \text{pentru } n = 0 \\ 0 & \text{pentru } n \neq 0 \end{cases}$.

În cazul aplicaţiei, funcţia pondere are expresia:

$h_n = T[\delta_n] = \alpha^n \cdot \sigma_n$, unde $\sigma_n = \begin{cases} 1 & \text{pentru } n \geq 0 \\ 0 & \text{pentru } n < 0 \end{cases}$

Prin urmare, **funcţia de transfer** a sistemului, ca transformată Fourier a funcţiei pondere, este dată de relaţia: $H(f) = \sum_{n=0}^{+\infty} h_n \cdot \exp(-j2\pi f \cdot n)$, iar rezultatul final se obţine din relaţia: $S_Y(f) = |H(f)|^2 \cdot S_X(f)$, unde:

$$S_X(f) = \mathrm{F}\{R_X(k)\} = \sum_{k=-\infty}^{+\infty} R_X(k) \cdot \exp(-j2\pi f \cdot k) = \sigma_X^2$$

deoarece, din ipoteză: $R_X(k) = \begin{cases} \sigma_X^2 & \text{pentru } k = 0 \\ 0 & \text{pentru } k \neq 0 \end{cases}$

* * *

În privinţa autocovarianţei sau, implicit, al coeficientului de corelaţie, aceasta/acesta arată cât de corelate pot fi valorile pe care un proces aleatoriu le poate lua în două momente diferite, respectiv doi paşi diferiţi. Mai mult, ţinând cont de expresia de calcul al coeficientului de corelaţie $\rho_X(n_2 - n_1) = \rho_X(k)$,

putem aprecia variaţia acesteia funcţie de "distanţa" k dintre cele două puncte de observare considerate, ale procesului. Astfel, în cazul proceselor discrete, acestea pot fi:

- **Procese necorelate**, numite şi **procese fără memorie**, dacă:

$$\rho_X(n_2 - n_1) = \rho_X(k) = 0, \ \forall k \geq 1 \qquad (9.3.13)$$

- **Procese corelate pe termen scurt SRD** (***S**hort **R**ange **D**ependent*), numite şi **procese cu memorie scurtă**, dacă:

$$\sum_k |\rho_X(k)| < \infty \qquad (9.3.14)$$

- **Procese corelate pe termen lung LRD** (***L**ong **R**ange **D**ependent*), numite şi **procese cu memorie lungă**, dacă:

$$\sum_k |\rho_X(k)| = \infty \qquad (9.3.15)$$

O altă mărime prin intermediul căreia se poate stabili gradul de corelare ce caracterizează un proces aleatoriu staţionar în sens larg, $\{X_n, n \geq 0\}$, este varianţa corespunzătoare procesului $\{X_j^{(m)}, j \geq 0\}$ obţinut prin medierea procesului original pe intervale (eşantioane) disjuncte, de dimensiune, adică:

$$X_j^{(m)} = \left(X_{j \cdot m - m + 1} + \ldots + X_{j \cdot m}\right) \cdot m^{-1} \qquad (9.3.16)$$

Ţinând cont de caracteristicile procesului original $\{X_n, n \geq 0\}$, varianţa procesului $\{X_j^{(m)}, j \geq 0\}$ este dată de relaţia:

$$\mathrm{VAR}\left[X^{(m)}\right] = \mathrm{VAR}[X] \cdot \left\{1 + 2\sum_{k=1}^{m} (1 - k/m) \cdot \rho_x(k)\right\} \cdot m^{-1} \qquad (9.3.17)$$

care, pentru valori mari ale lui m, se poate aproxima prin:

$$\mathrm{VAR}\left[X^{(m)}\right] = \mathrm{VAR}[X] \cdot \left\{1 + 2\sum_{k=1}^{m} \rho_x(k)\right\} \cdot m^{-1} \qquad (9.3.18)$$

Plecând de la relaţiile (9.3.17) şi (9.3.18) şi ţinând cont de condiţiile (9.3.13) - (9.3.15), se obţin următoarele condiţii echivalente:

- Un **proces** este **necorelat** dacă $\mathrm{VAR}\left[X^{(m)}\right]$ converge la zero proporţional cu valoarea inversă, m^{-1}, a dimensiunii eşantionului,
- Un **proces** este **corelat pe termen scurt** dacă $\mathrm{VAR}\left[X^{(m)}\right]$ converge la zero asimptotic proporţional cu m^{-1},
- Un **proces** este **corelat pe termen lung** dacă $\mathrm{VAR}\left[X^{(m)}\right]$ converge la zero mai încet decât m^{-1}.

O categorie tipică de proces corelat pe termen scurt este cel la care $\rho_X(k) \sim \alpha^{|k|}$, cu $0 < \alpha < 1$ şi $|k| \to \infty$, adică procesele pentru care coeficientul de corelaţie scade după o lege geometrică (exponenţială).

Notaţia $a(k) \sim b(k)$ pentru $|k| \to \infty$ înseamnă: $\lim_{|k|\to\infty} \frac{a(k)}{b(k)} = 1$, adică expresiile $a(k)$ şi $b(k)$ sunt **asimptotic proporţionale**.

Pentru procesele corelate pe termen lung, o categorie tipică este aceea la care $\rho_X(k) \sim |k|^{-\beta}$, cu $0 < \beta \le 1$ şi $|k| \to \infty$, adică procesele pentru care coeficientul de corelaţie scade după o lege hiperbolică.

Figura 9.3.1 prezintă drept exemplificare variaţia particulară a unor astfel de coeficienţi de corelaţie, împreună cu cazul necorelat.

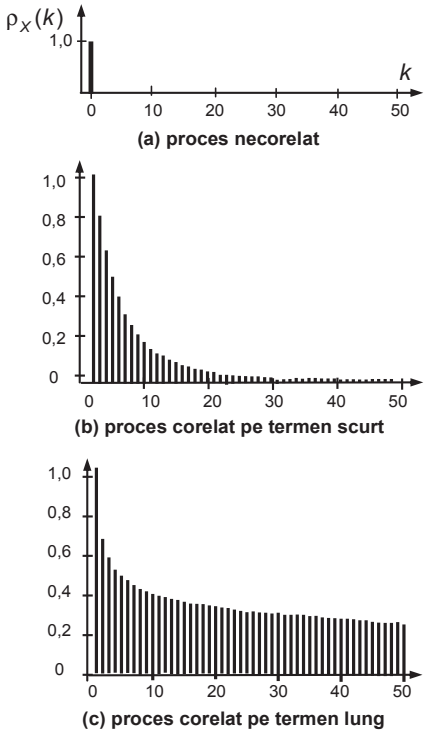


Figura 9.3.1: Exemple de variaţie a coeficienţilor de corelaţie

Aplicaţia 9.3.5

Luând în considerare condiţiile echivalente, precizate anterior, stabiliţi gradul de corelare al procesului:

i) secvenţă de variabile aleatorii i.i.d. după o lege Bernoulli de parametru p;

ii) secvenţă de variabile aleatorii i.i.d. după o lege exponenţială de parametru λ;

iii) secvenţa propusă în cadrul aplicaţiei 9.3.3.

Indicaţie: ii) a se vedea distribuţia *m*-Erlang.

9.4 Medii temporale şi condiţii de ergodicitate

Conceptele dezvoltate pe parcursul capitolelor anterioare au luat în considerare câteva modele teoretice de procese aleatorii pentru care s-au specificat direct sau indirect distribuţiile ce le caracterizează. Deseori, acest context nu este valabil în cazul proceselor aleatorii cu care ne confruntăm în realitate, fiind de cele mai multe ori imposibil de urmărit toate realizările atât din punct de vedere al numărului lor, cât şi al mulţimii de indexare (ambele, în general, infinite). În aceste condiţii, parametri proceselor urmărite (media, varianţa, autocorelaţia etc.) pot fi doar *estimaţi* pe baza unui ansamblu ***finit*** de măsurători (cu ajutorul cărora se poate, eventual, identifica şi modelul teoretic care descrie satisfăcător fenomenul studiat). Astfel, de exemplu media, $m_X(t)$, a unui proces aleatoriu, $X(t,\xi)$, se poate estima repetând experimentul şi aplicând următoarea formulă:

$$\hat{m}_X(t) = \sum_{i=1}^{n} X(t,\xi_i) \Big/ n \qquad (9.4.1)$$

unde n este numărul repetărilor, $X(t,\xi_i)$ este realizarea observată la repetarea numărul i şi $\hat{m}_X(t)$ *este o variabilă aleatorie* ale cărei realizări estimează mărimea căutată.

Aplicaţia 9.4.1

Estimaţi media unui proces aleatoriu, $X(t,\xi_i)$, care a generat prin repetare realizările din figura 9.4.1.

Rezolvare: observând realizările procesului precizat se obţine că valorile pentru $\hat{m}_X(t)$ sunt în ordine: $3/4$ pentru $0 \le t < 10$, $5/4$ pentru $10 \le t < 30$ etc.

* * *

În unele situaţii este avantajos să estimăm parametri unui proces aleatoriu prin interpretarea unei singure realizări, $X(t,\xi_i)$, utilizând, pentru definirea variabilei aleatorii care estimează media, următorul gen de relaţii:

- în cazul unui proces continuu:

$$\langle X(t)\rangle_\theta = 1/\theta \cdot \int_0^\theta X(t,\xi_i)\,\mathrm{d}t \qquad (9.4.2)$$

- în cazul unui proces discret:

$$\langle X_n\rangle_k = 1/k \cdot \sum_{i=1}^{k} X_i \qquad (9.4.3)$$

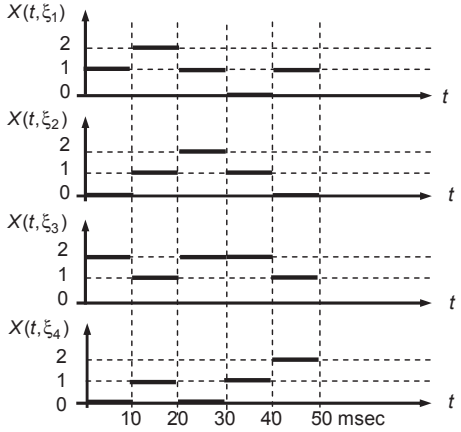


Figura 9.4.1: Patru realizări ale procesului propus în aplicaţia 9.4.1

Aplicaţia 9.4.2

Estimaţi media procesului aleatoriu, analizat în aplicaţia precedentă, prin medierea în timp a uneia din realizările sale, prezentate în figura 9.4.1.

* * *

Pentru a folosi formule de acest gen, este necesar să stabilim în ce condiţii *medierile în timp* (pe mulţimea *I* de indexare) converg (eventual către valorile aşteptate – parametri pe ansamblu) odată cu creşterea mulţimilor $[0,\theta]$ sau $\{1,2,\cdots,k\}$, luate în considerare. Aceste condiţii sunt specificate prin intermediul ***teoremelor de ergodicitate*** din categoria cărora face parte şi legea tare a numerelor mari care afirmă că dacă X_n este o secvenţă de variabile aleatorii (proces) independente şi identic distribuite de medie finită, $\mathrm{E}[X_n] = m$, atunci medierea în timp a unei realizări (eşantion al realizărilor variabilelor aleatorii indexate) converge cert (cu probabilitate 1) către media pe ansamblu:

$$P\left[\lim_{n\to\infty} 1/n \cdot \sum_{i=1}^{n} X_i = m\right] = 1 \qquad (9.4.4)$$

Mai mult, conform teoremei lui Doob-Birkhoff, pentru orice proces staţionar, variabilele aleatorii $\langle X(t)\rangle_\theta$, respectiv $\langle X_n\rangle_k$ converg aproape sigur către o constantă (care, conform exemplului următor, nu este însă în mod obligatoriu, media procesului în cauză !).

Aplicaţia 9.4.3

Fie procesul aleatoriu $X(t) = Y$, $\forall t \geq 0$ unde Y este o variabilă aleatorie de medie 0 şi spaţiul realizărilor $\mathrm{S}_Y = (-1,1)$.

i) Trasaţi pe acelaşi grafic câteva realizări particulare ale procesului;
ii) Precizaţi tipul de staţionaritate ce îl caracterizează;
iii) Comparaţi media pe ansamblu cu medierea unei singure realizări.

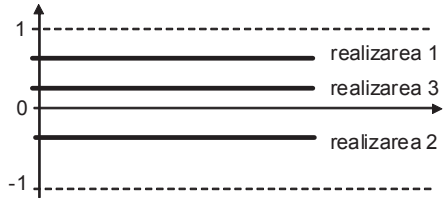


Figura 9.4.2: Trei realizări posibile ale procesului analizat în aplicaţia 9.4.2

Rezolvare: i) vezi figura 9.4.2;

iii) $m_X(t) = \mathrm{E}[X(t)] = \mathrm{E}[Y] = 0$ şi $\langle X(t)\rangle_\theta = 1/\theta \cdot \int_0^\theta y\,\mathrm{d}t = y \neq m_X(t)$,

unde y este o realizare particulară a variabilei aleatorii Y.

* * *

Domeniul relativ restrâns de aplicabilitate a teoremelor sus amintite impune identificarea unor clase mai largi de procese aleatorii pentru care medierile temporale tind către valorile aşteptate. O astfel de clasă este cea a proceselor staţionare în sens larg, $X(t)$, de medie constantă, $m_X(t) = m$ (medierile în timp oferind doar constante), care trebuie să respecte condiţia:

$$\lim_{\theta\to+\infty} \mathrm{VAR}\left[\langle X(t)_\theta\rangle\right] = \lim_{\theta\to+\infty} \mathrm{E}\left[\left(\langle X(t)_\theta\rangle - m\right)^2\right] = 0 \qquad (9.4.5)$$

pentru a putea considera că:

$$\lim_{\theta\to+\infty} \langle X(t)\rangle_\theta = m. \qquad (9.4.6)$$

Această condiţie, (9.4.5), cu formă analogă în cazul proceselor discrete, impune, de fapt, ca la limită, când $\theta \to +\infty$, variabila aleatorie $\langle X(t)\rangle_\theta$ să aibă o varianţă neglijabilă, astfel încât toate realizările sale să se găsească în vecinătatea imediată a aceleiaşi valori (vezi figura 9.4.3).

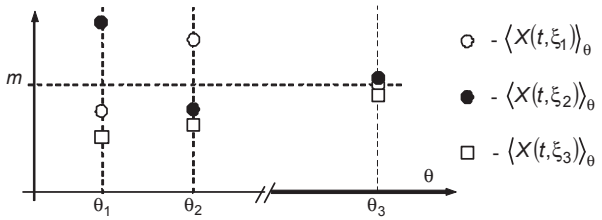


Figura 9.4.3: Convergenţă asimptotică a realizărilor – exemple

În practică, verificarea condiţiei (9.4.5) se poate face recurgând la următoarea relaţie echivalentă [Gar89]:

$$\lim_{\theta\to+\infty} 1/\theta\cdot\int_0^{\theta}(1-u/\theta)\cdot C_X(u)\,\mathrm{d}u = 0 \qquad (9.4.7)$$

unde $C_X(u)$ este autocovarianţa procesului analizat.

Demonstraţie: Fie un proces staţionar în sens larg. Rezultatul medierii în timp a unei singure realizări limitate la intervalul $[0,\theta]$ este o variabilă aleatorie a cărei medie este:

$$\mathrm{E}\left[\langle X(t)\rangle_\theta\right] = \mathrm{E}\left[1/\theta\cdot\int_0^{\theta}X(t)\,\mathrm{d}t\right] = 1/\theta\cdot\int_0^{\theta}\mathrm{E}[X(t)]\,\mathrm{d}t = m \qquad (9.4.8)$$

Conform acestui rezultat, variabila aleatorie $\langle X(t)\rangle_\theta$ este un **estimator nedeplasat** (**unbiased**) al mărimii căutate, $\mathrm{E}[X(t)] = m$.

În privinţa varianţei acestei variabile aleatorii, relaţia sa de calcul este:

$$\begin{aligned}\mathrm{VAR}\left[\langle X(t)\rangle_\theta\right] &= \mathrm{E}\left[\left(\langle X(t)\rangle_\theta - m\right)^2\right]\\ &= \mathrm{E}\left[\left\{1/\theta\cdot\int_0^{\theta}(X(t)-m)\,\mathrm{d}t\right\}\cdot\left\{1/\theta\cdot\int_0^{\theta}(X(t')-m)\,\mathrm{d}t'\right\}\right]\\ &= 1/\theta^2\cdot\int_0^{\theta}\int_0^{\theta}\mathrm{E}\left[(X(t)-m)\cdot(X(t')-m)\right]\mathrm{d}t\,\mathrm{d}t'\\ &= 1/\theta^2\cdot\int_0^{\theta}\int_0^{\theta}C_X(t,t')\,\mathrm{d}t\,\mathrm{d}t'\end{aligned} \qquad (9.4.9)$$

Ţinând cont că în cazul unor procese staţionare în sens larg $C_X(t,t') = C_X(t-t')$, ultima egalitate devine:

$$\mathrm{VAR}\left[\langle X(t)\rangle_\theta\right] = 1/\theta^2\cdot\int_0^{\theta}\int_0^{\theta}C_X(t-t')\,\mathrm{d}t\,\mathrm{d}t' \qquad (9.4.10)$$

Pentru a rezolva integrala, este util de observat că integrantul rămâne constant pe oricare din segmentele generate de liniile de ecuaţie $t-t'=u$, conţinute în regiunea de integrare prezentată în figura 9.4.4.

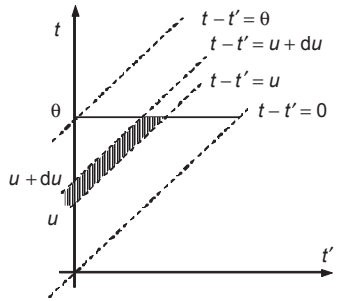


Figura 9.4.4: Domeniu de integrare pentru integrala din ecuaţia (9.4.10)

În aceste condiţii, integrala (volumul total al corpului delimitat de suprafaţa $C_X(t-t')$ în dreptul domeniului de integrare) poate fi calculată ca o sumă de volume infinitezimale obţinute prin înmulţirea unor arii de genul celei haşurate în figura 9.4.4 cu *înălţimea* corespunzătoare, obţinându-se relaţia căutată:

$$\mathrm{VAR}\left[\langle X(t)\rangle_\theta\right] = \frac{1}{\theta^2}\cdot\int_0^{\theta}\underbrace{C_X(u)}_{\text{înălţimea}}\cdot\underbrace{(\theta-u)\,\mathrm{d}u}_{\text{aria suprafeţei haşurate}} = \frac{1}{\theta}\cdot\int_0^{\theta}C_X(u)\cdot\left(1-\frac{u}{\theta}\right)\mathrm{d}u \qquad (9.4.11)$$

Condiţia de ergodicitate (9.4.7) este valabilă, desigur, şi pentru procesele aleatorii discrete staţionare în sens larg, când se prezintă în forma următoare:

$$\lim_{k\to\infty}\frac{1}{k}\cdot\sum_{i=1}^{k}\left(1-\frac{i}{k}\right)\cdot C_X(i) = 0 \qquad (9.4.12)$$

Aplicaţia 9.4.4

Verificaţi dacă este ergodic un proces caracterizat prin funcţia de autocovarianţa $C_X(\tau) = \mathrm{e}^{-\alpha\cdot\tau},\ \tau\ge 0$

Rezolvare : $\mathrm{VAR}\left[\langle X(t)\rangle_\theta\right] = \frac{1}{\theta}\cdot\int_0^{\theta}\mathrm{e}^{-\alpha\cdot u}\cdot\left(1-\frac{1}{\theta}\right)\mathrm{d}u < \frac{1}{\theta}\int_0^{\theta}\mathrm{e}^{-\alpha\cdot u}\mathrm{d}u = \frac{1-\mathrm{e}^{-\alpha\cdot\theta}}{\alpha\cdot\theta}\underset{\theta\to\infty}{\to}0$

* * *

Condiţia de ergodicitate, stabilită anterior, modificată în sensul că limita trebuie să tindă către o valoare constantă ($<\infty$), nu obligatoriu 0, este aplicabilă şi în cazul proceselor aleatorii asimptotic staţionare în sens larg. Un ***proces*** este ***asimptotic staţionar în sens larg*** dacă [Mac87]:

$$\lim_{t\to\infty} m_X(t) = m \qquad (9.4.13)$$

$$\lim_{t\to\infty} C_X(t, t+\tau) = C_X(\tau) \qquad (9.4.14)$$

Conceptele şi relaţiile anterioare, cuprinse în acest capitol au fost introduse luând în considerare doar procese aleatorii de tip cauzal, care se desfăşoară dintr-un anumit moment (considerat, în general, originea axei timpului). Cu toate acestea, ele privesc şi procesele aleatorii necauzale, singura modificare ce trebuie făcută la nivel de relaţii fiind cea a schimbării intervalului de observare a unei realizări din $(0,\theta)$, respectiv $\{1,\cdots,k\}$ în $(-\theta,+\theta)$, respectiv $\{-k,\cdots,0,\cdots,+k\}$. În acest fel, pentru definiţia mediei unei realizări se obţine:

$$\langle X(t)\rangle_{\theta} = (2\theta)^{-1}\cdot\int_{-\theta}^{+\theta} X(t,\xi)\,\mathrm{d}t \qquad (9.4.15)$$

iar pentru condiţia de ergodicitate:

$$\lim_{\theta\to\infty}\frac{1}{2\theta}\int_{-2\theta}^{2\theta}\left(1-\frac{|u|}{2\theta}\right)\cdot C_X(u)\,\mathrm{d}u = 0 \qquad (9.4.16)$$

şi respectiv:

$$\lim_{k\to\infty}\frac{1}{2k+1}\cdot\sum_{i=-k}^{+k}\left(1-\frac{|i|}{k}\right)\cdot C_X(i) = 0 \qquad (9.4.17)$$

În încheierea acestui subcapitol este util de observat că staţionaritatea unui proces aleatoriu este deseori dificil de demonstrat experimental. Din acest motiv, în general, ne bazăm pe intuiţie atunci când considerăm ipoteza staţionarităţii, rămânând ca ulterior să verificăm veridicitatea sa.

9.5 Procese autosimilare

Un proces $Y(t)$ prezintă **autosimilitudine**, altfel spus, este **autosimilar SS** (***SelfSimilar***), dacă există un număr unic H, $H>0$, denumit **parametru Hurst**, pentru care:

$$\{X(a\cdot t)\}\overset{d}{=}\{a^H\cdot X(t)\},\ \forall a>0 \text{ şi } t\geq 0 \qquad (9.5.1)$$

în care notaţia $\overset{d}{=}$ înseamnă egalitate în ceea ce privesc toate distribuţiile de ordin finit ale celor două procese asociate.

Aplicaţia 9.5.1

Luând în considerare modurile de specificare ale unui proces aleatoriu, prezente în (9.2.1)-(9.2.3), detaliaţi definiţia dată similarităţii, din relaţia (9.5.1).

* * *

Aplicaţia 9.5.2

Verificaţi dacă secvenţele de numere aleatorii i.i.d. sunt procese autosimilare.

* * *

Denumirea parametrului referit în definiţia unui proces autosimilar vine de la numele inginerului englez H.E. Hurst care a descoperit primul această proprietate, cu ocazia studiilor sale întreprinse în vederea proiectării barajului de la Assuan, din Egipt (anume pentru stabilirea capacităţilor de acumulare). La modul principial, obiectivul urmărit în cadrul acestui gen de proiect este ca, în condiţiile menţinerii unui debit de ieşire constant, egal cu debitul mediu de intrare, fluctuaţiile cantităţii de apă din lacul de acumulare să nu ducă la golirea sau revărsarea acestuia. Întrucât aceste fluctuaţii depind de precipitaţiile locale, Hurst a întreprins o serie de studii statistice ale datelor hidrologice disponibile, rezultatul obţinut, valabil şi în cazul altor domenii, precum cel al reţelelor de comunicaţii [LeL93] precizând că: sunt zile secetoase şi zile ploioase, sunt săptămâni secetoase şi săptămâni ploioase, sunt luni secetoase şi luni ploioase, sunt ani secetoşi ...., decenii secetoase etc. Cu alte cuvinte, la orice scară temporală am studia fenomenul, comportamentul acestuia este asemănător.

Un exemplu, ilustrat în figura 9.5.1, îl oferă şi procesul real ce a fost "măsurat" în decurs de o oră, [LeL93], alături de un proces Poisson de aceeaşi medie şi deviaţie standard cu procesul real. Imaginile din figură, considerate pe verticală de sus în jos, oferă rezultatele măsurătorilor efectuate în 60 de intervale distincte, ce acoperă perioade de timp cu durată în creştere, şi anume de 6, 60, 600 şi 3600 secunde.

Se poate observa că, pe măsură ce scara timpului creşte, pentru procesul Poisson fluctuaţiile aproape au dispărut, iar pentru procesul real scara timpului nu are o influenţă notabilă, ceea ce implică faptul că un asemenea proces este autosimilar. Desigur că, valorile măsurate diferă semnificativ ca plajă a realizărilor posibile funcţie de intervalele de timp considerate. Prin urmare, în vederea "comparării" rezultatelor ce iau în considerare intervale diferite, precum t (de exemplu, un minut) şi a (de exemplu, o oră) este necesară introducerea unui factor de ajustare, nu altul decât termenul a^H din definiţia (9.5.1).

Justificarea diferenţei evidenţiate în figura 9.5.1, dintre un proces Poisson şi un proces autosimilar, pleacă de la considera intervalul minim pe care se face contorizarea, egal cu 1, dimensiunea avută în vedere, în cazul unui experiment real putând fi $\mu\mathrm{sec}$, msec etc. În aceste condiţii, notând cu $X_P(t)$ procesul Poisson şi cu $X_{SS}(t)$ procesul autosimilar, avem următoarele rezultate:

$$m_{X_P(1)}=\lambda \text{ şi } \sigma_{X_P(1)}=\sqrt{\lambda};\ m_{X_P(t)}=t\cdot m_{X_P(1)} \text{ şi } \sigma_{X_P(t)}=\sqrt{t}\cdot\sigma_{X_P(1)};$$

$$m_{X_{SS}(t)}=t^H\cdot m_{X_{SS}(1)} \text{ şi } \sigma_{X_{SS}(t)}=t^H\cdot\sigma_{X_{SS}(1)}$$

Ultimele două egalităţi se obţin făcând substituţiile: $t\leftarrow 1$ şi $a\leftarrow t$ în relaţia 9.5.1. Prin urmare, în cazul procesului Poisson, coeficientul de variaţie, C_X, este proporţional cu $t^{-1/2}$, ceea ce face ca procesul incrementărilor pe eşantioane succesive, de durată t, să devină tot mai ”neted” odată cu creşterea acestuia. În schimb, în cazul proceselor autosimilare, coeficientul de variaţie

rămâne constant (egal cu 1) ceea ce face ca procesul însumării pe eşantioane succesive, de durată t, să-şi păstreze forma, indiferent de mărimea acestora.

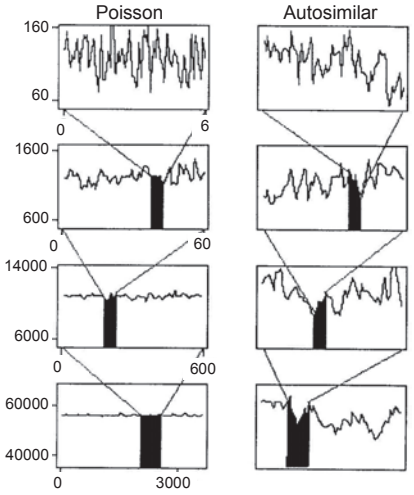


Figura 9.5.1: Exemplu de proces autosimilar comparat cu un proces Poisson, cu acelaşi debit mediu

Discuţia de mai sus evidenţiază faptul că un proces autosimilar nu este staţionar fiind puţin interesant din punct de vedere practic atât timp cât nu este caracterizat de proprietăţi care să-i confere predictibilitate. În consecinţă, în cele ce urmează vom considera **procese**le **autosimilare cu exponent *H* şi incrementare staţionară**, notate pe scurt ***H*-sssi** (***H*-*s*elf *s*imilar *s*tationary *i*ncrement**). Legătura dintre un proces H-sssi discret, X_n, şi incrementările $\varepsilon_n^{(m)}$, pe intervale disjuncte, alăturate şi de lungime m paşi, cu exemplificare în figura 9.5.2, este dată de relaţia:

$$\varepsilon_n^{(m)} = X_{m\cdot n+m} - X_{m\cdot n} \qquad (9.5.2)$$

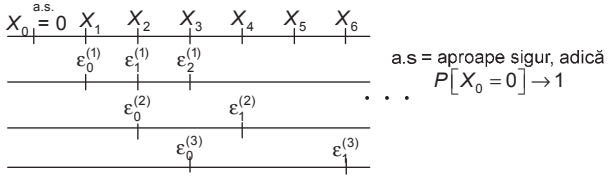


Figura 9.5.2: Proces discret şi incrementări asociate

Plecând de la relaţia (9.5.2), coeficientul de corelaţie al procesului de incrementare pe diverse lungimi de intervale, $\varepsilon_n^{(m)}$, se obţine ţinând cont de următoarea proprietate privitoare la autocorelaţia proceselor H-sssi:

$$\mathrm{E}\left[X(t)\cdot X(s)\right] = \frac{1}{2}\cdot\left\{t^{2H} + s^{2H} - |t-s|^{2H}\right\}\cdot \mathrm{E}\left[X^2(1)\right] \qquad (9.5.3)$$

cu condiţia $\mathrm{E}\left[X^2(1)\right] < \infty$, care implică, de asemenea: $0 < H \le 1$.

Aplicaţia 9.5.3

Verificaţi corectitudinea relaţiei (9.5.3).

Indicaţie: se dezvoltă $\mathrm{E}\left[\left(X(t)-X(s)\right)^2\right]$.

* * *

În vederea calculării coeficientului de corelaţie, ipoteza de lucru este: **incrementările sunt de medie zero**, $\mathrm{E}\left[\varepsilon_n^{(m)}\right] = 0$. În consecinţă:

$$\begin{aligned} C_{\varepsilon^{(m)}}(n) = \mathrm{E}\left[\varepsilon_0^{(m)}\cdot\varepsilon_n^{(m)}\right] &= \mathrm{E}\left[\left\{X_m - X_0\right\}\cdot\left\{X_{m\cdot n+m} - X_{m\cdot n}\right\}\right] \\ &\overset{X_0=0}{=} \mathrm{E}\left[X_m\cdot X_{m\cdot n+m}\right] - \mathrm{E}\left[X_m\cdot X_{m\cdot n}\right] \\ &\overset{(6.5.3)}{=} \frac{1}{2}\cdot\left\{(n+1)^{2H} - 2n^{2H} + (n-1)^{2H}\right\}\cdot\mathrm{E}\left[\left(m^H\cdot X_1\right)^2\right] \end{aligned} \qquad (9.5.4)$$

Deci, ţinând cont că:

$\mathrm{E}\left[\left(m^H\cdot X_1\right)^2\right] = \mathrm{E}\left[{X_m}^2\right]$, procesul fiind autosimilar (vezi relaţia (9.5.1),

şi $\mathrm{VAR}\left[\varepsilon_0^{(m)}\right] = \mathrm{VAR}\left[X_m\right] = \mathrm{E}\left[{X_m}^2\right]$, respectiv:

$$\mathrm{VAR}\left[\varepsilon_n^{(m)}\right] = \mathrm{VAR}\left[X_{m\cdot n+m} - X_{m\cdot n}\right] \overset{\text{s.i.}}{=} \mathrm{VAR}\left[X_m\right],$$

deoarece incrementările sunt staţionare (**s**.**i**.), însemnă că, pentru coeficientul de corelaţie, este valabilă următoarea relaţie de calcul:

$$\rho_{\varepsilon^{(m)}}(n) = \rho(n) = 1/2\cdot\left\{(n+1)^{2H} - 2n^{2H} + (n-1)^{2H}\right\} \qquad (9.5.5)$$

La limită, când $n \to \infty$, relaţia (9.5.5) poate fi pusă sub forma:

$$\rho(n)\begin{cases} \sim H(2H-1)\,n^{2(H-1)} & \text{pentru } H \ne 1/2 \\ = 0 & \text{pentru } H = 1/2 \text{ şi } n \ge 1 \end{cases} \qquad (9.5.6)$$

care conduce la următoarele concluzii, privind impactul parametrului Hurst asupra coeficientului de corelaţie:

- dacă $0 < H < 1/2 \rightarrow$ corelaţia este negativă, $\rho(n) < 0$, pentru $n \geq 1$, şi $\sum_n |\rho(n)| < \infty$, adică dependenţa este pe termen scurt,
- dacă $H = 1/2 \rightarrow$ corelaţia este nulă, $\rho(n) = 0$ pentru $n \geq 1$, adică incrementările $\varepsilon_n^{(m)}$ sunt necorelate,
- dacă $1/2 < H < 1 \rightarrow$ corelaţia este pozitivă $\rho(n) > 0$, pentru $n \geq 1$, şi $\sum_n |\rho(n)| = \infty$, adică dependenţa este pe termen lung
- dacă $H = 1 \rightarrow$ corelaţia $\rho(n) = 1$, $n \geq 0$, adică un proces degenerat, fără utilitate practică.

Relaţia (9.5.5), şi respectiv (9.5.6), precizează faptul că indiferent de gradul de agregare temporală (numărul m de paşi considerat) coeficientul de corelaţie rămâne neschimbat. În consecinţă, ţinând cont că **staţionaritatea în sens larg** mai poartă şi denumirile: **staţionaritate slabă** (*weak*), **staţionaritatea covarianţei** sau **staţionaritatea de ordinul doi**, sunt adevărate următoarele două enunţuri:

- un proces $X(t)$ este **exact H-sssi**, adică **autosimilar de ordinul doi cu parametru H**, dacă este verificată relaţia (9.5.5),
- un proces $X(t)$ este **asimptotic H-sssi**, adică **asimptotic autosimilar de ordinul doi cu parametru H**, dacă:

$$\lim_{m \to \infty} \rho_{\varepsilon^{(m)}}(n) = \frac{1}{2}\left\{(n+1)^{2H} - 2n^{2H} + (n-1)^{2H}\right\} \qquad (9.5.7)$$

În denumirile de mai sus:

- ”de ordinul doi” precizează faptul că valoarea coeficientului de corelaţie a incrementărilor depinde doar de ecartul (n) dintre eşantioane,
- ”autosimilar” precizează că valoarea coeficientului de corelaţie a incrementărilor nu depinde de dimensiunea eşantionului,
- ”asimptotic” precizează că, de la un anumit m ($m \to \infty$), coeficientul de corelaţie a incrementărilor nu mai depinde de dimensiunea eşantionului.

Aplicaţia 9.5.4

Se consideră **mişcare browniană** (standard) acel proces, $\{B(t), t \geq 0\}$, care satisface următoarele proprietăţi:

- $B(0) = 0$ a.s.,
- incrementările sunt staţionare şi independente,
- pentru orice $t > 0$, $B(t)$ urmează o distribuţie Gauss de medie zero şi varianţă $\sigma^2 \cdot t$,
- realizările sale sunt a.s. continue.

Arătaţi că mişcarea browniană este un proces $1/2$-sssi.

Indicaţie: se arată că pentru orice $a > 0$, $\left\{a^{-1/2} B(a \cdot t), t \geq 0\right\}$ este tot o mişcare browniană.

9.6 Procese aleatorii tip sumă

Adunarea unei secvenţe de variabile aleatorii i.i.d., $X_1, X_2, \cdots$, generează ***procese aleatorii discrete de tip sumă*** conform relaţiei:

$$\begin{aligned} S_n &= X_1 + X_2 + \cdots + X_n \\ &= S_{n-1} + X_n \end{aligned}, \text{ cu } n = 1,2,\cdots \text{ şi } S_0 = 0 \qquad (9.6.1)$$

A doua egalitate pune în evidenţă faptul că procesele sumă deţin memorie, valoarea s_n pe care procesul o poate lua la un anumit pas (moment), n, depinzând de valoarea s_{n-1}, luată cu un pas în urmă (momentul precedent).

Formula de recurenţă cuprinsă în relaţia (9.6.1) se poate reprezenta grafic, ca în figura 9.6.1(a). Prin comparare cu figura 9.6.1(b), se evidenţiază faptul că procesul aleatoriu de numărare este un caz particular de ***proces aleatoriu autoregresiv*** de ordinul unu AR(1), [Gar89], [MDS01], al cărui caz general AR(p) urmează a fi tratat în secţiunea 9.7.

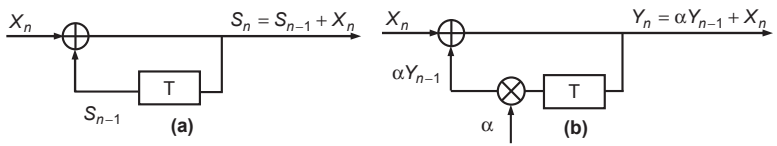


Figura 9.6.1: Schema de generare a proceselor aleatorii: (a) de tip sumă; (b) autoregresive de ordinul unu

Aplicaţia 9.6.1

Determinaţi media şi varianţa procesului binomial de numărare.

Rezolvare: **Procesul binomial de numărare** este un proces sumă dat de relaţia: $S_n = I_1 + I_2 + \ldots + I_n$, cu $n \geq 1$, unde $I_1, I_2, \cdots$ este o secvenţă de variabile aleatorii i.i.d. care au spaţiul realizărilor $\{0,1\}$ şi urmează o lege Bernoulli de parametru $p = P\left[I_j = 1\right]$, $\forall j$. În acest caz, probabilităţile elementare de ordinul 1, care specifică acest proces, precum şi mărimile căutate: media şi varianţa corespund unei variabile aleatorii binomiale de parametru n şi p, exprimate ca:

- $P[S_n = j] = \begin{cases} C_n^j \cdot p^j \cdot (1-p)^{n-j} & \text{pentru } 0 \leq j \leq n \\ 0 & \text{în rest} \end{cases}$;

- $m_S(n) = n \cdot p$;
- $\mathrm{VAR}[S_n] = n \cdot p \cdot (1-p)$.

* * *

Procesele aleatorii discrete de tip sumă prezintă proprietăţile de **incrementare independentă** pe intervale disjuncte şi **incrementare staţionară** pe intervale de aceeaşi lungime, indiferent de unde încep acestea. Prima proprietate specifică faptul că incrementările pe intervale disjuncte sunt variabile aleatorii independente, iar a doua că incrementările pe intervale egale urmează, ca variabile aleatorii, aceeaşi distribuţie.

Aplicaţia 9.6.2

Determinaţi expresia distribuţiei de ordinul k a unui proces sumă, S_n, funcţie de distribuţiile corespunzătoare indexului iniţial, n_1, şi incrementărilor pe diverse intervale, care despart ceilalţi indecşi luaţi în considerare, $n_2, \cdots, n_k$. Variabile aleatorii i.i.d. care alcătuiesc secvenţa însumată sunt:

(a) discrete (caz particular: distribuţia binomială de parametru p, $k = 2$);

(b) continue (caz particular: distribuţia exponenţială de parametru λ).

Rezolvare: (a) În acest caz, procesul este evaluat discret, putând fi descris de probabilităţile elementare de legătură de ordinul *k* care se pot exprima funcţie de distribuţiile menţionate în textul aplicaţiei, precum urmează:

$$P\left[S_{n_1} = x_1 \,\&\, S_{n_2} = x_2 \,\&\cdots\&\, S_{n_k} = x_k\right]$$

$$= P\left[S_{n_1} = x_1 \,\&\, S_{n_2} - S_{n_1} = x_2 - x_1 \,\&\cdots\&\, S_{n_k} - S_{n_{k-1}} = x_k - x_{k-1}\right]$$

$$\overset{\text{independenţa incrementării}}{=} P\left[S_{n_2} = x_{n_1}\right] \cdot P\left[S_{n_2} - S_{n_1} = x_2 - x_1\right] \cdot \ldots \cdot P\left[S_{n_k} - S_{n_{k-1}} = x_k - x_{k-1}\right] \quad (9.6.2)$$

$$\overset{\text{independenţa incrementării}}{=} P\left[S_{n_1} = x_1\right] \cdot P\left[S_{n_2 - n_1} = x_2 - x_1\right] \cdot \ldots \cdot P\left[S_{n_k - n_{k-1}} = x_k - x_{k-1}\right]$$

Dacă se particularizează această expresie pentru cazul particular propus, atunci se obţine relaţia următoare:

$$P\left[S_{n_1} = x_1 \,\&\, S_{n_2} = x_2\right] = P\left[S_{n_1} = x_1\right] \cdot P\left[S_{n_2 - n_1} = x_2 - x_1\right]$$

$$= C_{n_2 - n_1}^{x_2 - x_1} \cdot C_{n_1}^{x_1} \cdot p^{x_2} \cdot (1-p)^{n_2 - x_2}$$

(b) În mod similar, pentru secvenţa continuă, se obţine:

$$f_{S_{n_1}, S_{n_2}, \cdots, S_{n_k}}(x_1, x_2, \cdots, x_k) = f_{S_{n_1}}(x_1) \cdot f_{S_{n_2 - n_1}}(x_2 - x_1) \cdot \ldots \cdot f_{S_{n_k - n_{k-1}}}(x_k - x_{k-1}) \quad (9.6.3)$$

* * *

Un exemplu de proces aleatoriu continuu de tip sumă, care va fi deseori referit pe parcursul acestei lucrări, îl constituie ***procesul Poisson de numărare.*** O realizare posibilă a acestui proces poate fi cea din figura 9.6.2, unde momentele $S_1, S_2, \cdots$ pot reprezenta, de exemplu, momentele de sosire a unor clienţi într-un sistem de servire.

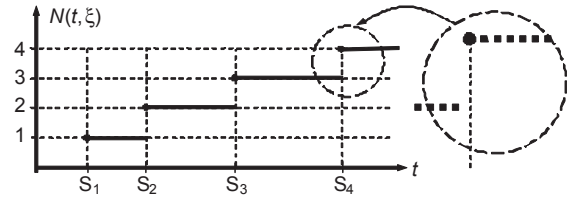


Figura 9.6.2: Exemplu de realizare a unui proces Poisson şi aproximarea sa printr-un proces binomial

Procesul $N(t)$, astfel considerat, reprezintă deci numărul de apariţii ale unui anumit eveniment (sosirea unui client în cazul exemplului menţionat anterior), în intervalul $[0,t]$. Fiind un proces continuu, specificarea sa se poate face aplicând metoda trecerii la limita infinitezimală a unui proces discret echivalent. În acest sens, intervalul $[0,t]$ se împarte în n subintervale egale, de durată $\delta = t/n$, şi se fac următoarele supoziţii:

1. Probabilitatea ca într-un subinterval δ evenimentul considerat să apară de două sau de mai multe ori este foarte mică, neglijabilă în comparaţie cu probabilitatea altor cazuri (apariţia o singură dată sau de loc).

2. Apariţia sau nu a evenimentului într-un interval este independentă de realizările observate în intervale disjuncte acestuia.

Prin asumarea acestor ipoteze, procesul analizat se aproximează cu un proces de numărare binomial, în fiecare subinterval δ desfăşurându-se câte un experiment de tip Bernoulli independent de precursorul său. Considerând că, din punct de vedere al procesului continuu, numărul realizărilor într-un anumit interval de timp (de exemplu 1 secundă) este în medie de λ apariţii/sec (sau incrementări/sec), probabilitatea p a succesului ce caracterizează procesul aproximativ de numărare binomial se determină din relaţia:

$$n \cdot p = \lambda \cdot t \quad (9.6.4)$$

Aceasta reprezintă numărul mediu de sosiri în intervalul $[0,t]$, exprimat în primul membru din punct de vedere al procesului binomial, respectiv din cel al procesului continuu de numărare în al doilea membru. În aceste condiţii, atunci când $n \to \infty$, respectiv $\delta = t/n \to 0$, şi $p \to 0$, pentru ca produsul $n \cdot p$ din

relaţia (9.6.4) să fie finit, procesul discret de numărare se transformă într-un proces continuu de numărare, a cărui distribuţie de ordinul 1 este de tip Poisson (de unde şi denumirea procesului), de parametru $\alpha = \lambda \cdot t$:

$$P[N(t)=k] = (\lambda \cdot t)^k / k! \cdot \exp(-\lambda \cdot t), \text{ pentru } k \geq 0 \qquad (9.6.5)$$

Aplicaţia 9.6.3

Stabiliţi dacă procesele de numărare, binomial şi Poisson, sunt autosimilare.

Indicaţie: se fac verificări la nivel de medie şi varianţă.

* * *

Fiind derivat dintr-un proces discret de tip sumă de variabile aleatorii i.i.d., procesul Poisson ***moşteneşte*** proprietăţile de incrementare independentă şi incrementare staţionară, fapt ce permite scrierea oricărei distribuţii de legătură conform regulii următoare, ce este precizată în cazul considerării doar a două momente t_1 şi t_2 $(t_1 < t_2)$:

$$\begin{aligned} P\left[\{N(t_1)=i\} \& \{N(t_2)=j\}\right] &= P[N(t_1)=i] \cdot P[N(t_2)-N(t_1)=j-i] \\ &= P[N(t_1)=i] \cdot P[N(t_2-t_1)=j-i] \qquad (9.6.6) \\ &= \frac{(\lambda t_1)^i}{i!} \cdot e^{-\lambda t_1} \cdot \frac{\{\lambda (t_2-t_1)\}^{j-i}}{(j-i)!} \cdot e^{-\lambda (t_2-t_1)} \end{aligned}$$

Aplicaţia 9.6.4

Sosirile clienţilor într-un sistem de servire urmează un proces Poisson de rată medie 15 sosiri/min. Determinaţi probabilitatea ca pe parcursul unui minut să sosească 2 clienţi în primele 10 secunde şi 3 clienţi în ultimele 10 secunde.

* * *

Un rezultat important, legat de procesele Poisson, este faptul că timpii dintre două sosiri succesive, T_i, sunt variabile aleatorii i.i.d. după o lege exponenţială de medie $1/\lambda$. Demonstraţia acestei afirmaţii se bazează tot pe discretizarea timpului în intervale (δ) ale căror lungimi devin infinitezimale $(\delta \to 0)$. Astfel, [Gar89], probabilitatea ca timpul T între două sosiri succesive să depăşească t secunde se aproximează iniţial cu probabilitatea ca în cele n subintervale, ce acoperă durata t, să nu apară nici o sosire, adică:

$$P[T > t] \cong (1-p)^n \text{ cu } n = \lfloor t/\delta \rfloor,$$

unde simbolul $\lfloor * \rfloor$ corespunde operaţiei de trunchiere la întreg. Trecând la limită, ($n \to \infty$, adică $\delta \to 0$), se obţine expresia distribuţiei căutate:

$$P[T > t] = \lim_{n \to \infty} (1 - \lambda \cdot t/n)^n = \exp(-\lambda \cdot t) \qquad (9.6.7)$$

În privinţa independenţei variabilelor T_i, aceasta este asigurată de independenţa variabilelor aleatorii geometrice ce corespund timpilor dintre două

sosiri succesive în cazul procesului discret folosit în faza de aproximare. Cunoscând distribuţia timpilor dintre două sosiri succesive, se pot determina caracteristicile probabilistice ale timpului până la apariţia celui de al n-lea eveniment (sosire). Acest timp este tot un proces sumă dat de relaţia:

$$S_n = T_1 + T_2 + \ldots + T_n$$

unde T_j, $j = \overline{1,n}$ reprezintă timpii între două sosiri succesive. Conform aplicaţiei 3.3.3.3, o astfel de sumă de variabile aleatorii urmează o distribuţie de tip n-Erlang de forma:

$$f_{S_n}(x) = \frac{(\lambda \cdot x)^{n-1}}{(n-1)!} \cdot \lambda \cdot \exp(-\lambda \cdot x) \text{ pentru } x \geq 0 \qquad (9.6.8)$$

Aplicaţia 9.6.5

Determinaţi media şi varianţa timpului până la sosirea a 10 clienţi într-un sistem ştiind că sosirea urmează un proces Poisson de medie 20 sosiri/min.

Indicaţie: Se pot utiliza relaţiile oferite de distribuţia m-Erlang sau se poate pleca de la faptul că timpii dintre două sosirii succesive sunt variabile aleatorii i.i.d., ceea ce permite scrierea relaţiilor: $E[S_{10}] = 10 \cdot E[T]$ şi $VAR[S_{10}] = 10 \cdot VAR[T]$, unde $E[T]$ şi $VAR[T]$ reprezintă parametrii distribuţiei exponenţiale *ataşate* timpului dintre două sosiri succesive.

* * *

Aplicaţia 9.6.6

Într-un sistem de servire, clienţii sosesc după un proces Poisson de rată medie λ. Presupunând că în intervalul $[0,\theta]$ apare un singur client, să se determine distribuţia momentului sosirii, X, a respectivului client, raportat la extremitatea inferioară (0) a intervalului.

Rezolvare: Presupunerea menţionată în textul aplicaţiei corespunde evenimentului $\{N(\theta)=1\}$, unde $N(t)$ este procesul Poisson care guvernează sosirea clienţilor în sistem. În aceste *condiţii*, distribuţia căutată este dată de relaţia: $P[X \leq x] = P[N(x)=1 \mid N(\theta)=1]$. Prin dezvoltarea expresiei pentru probabilitatea condiţionată, relaţia anterioară conduce la următorul rezultat:

$$\begin{aligned} P[X \leq x] &= \frac{P[\{N(x)=1\} \& \{N(\theta)=1\}]}{P[N(\theta)=1]} = \frac{P[\{N(x)=1\} \& \{N(\theta)-N(x)=0\}]}{P[N(\theta)=1]} \\ &= \frac{P[\{N(x)=1\} \& \{N(\theta-x)=0\}]}{P[N(\theta)=1]} = \frac{P[N(x)=1] \cdot P[N(\theta-x)=0]}{P[N(\theta)=1]} \\ &= \frac{\lambda x \cdot e^{-\lambda x} \cdot e^{-\lambda(\theta-x)}}{\lambda \theta \cdot e^{-\lambda \cdot \theta}} = \frac{x}{\theta} \end{aligned}$$

Expresia obţinută semnalează faptul că momentul X al sosirii unui client în intervalul $[0,\theta]$ urmează o distribuţie uniformă. Acest rezultat justifică de ce sunt echivalente exprimările: *clienţii sosesc la întâmplare în intervalul considerat* şi *clienţii sosesc conform unui proces Poisson*.

* * *

Privită din punctul de vedere al timpului dintre două sosiri (incrementări) succesive, rata de sosire, λ, ce corespunde unui proces Poisson este dată de inversa mediei timpului respectiv, adică:

$$\lambda = 1/\mathrm{E}[T] \qquad (9.6.9)$$

Această relaţie, caracteristică proceselor Poisson, este valabilă şi într-un cadru mai larg, şi anume al ***proceselor de numărare a reînnoirilor***, la care timpul dintre două sosiri (incrementări) succesive respectă o distribuţie arbitrară (nu doar exponenţială). Urmează demonstraţia acestei afirmaţii [Gar89].

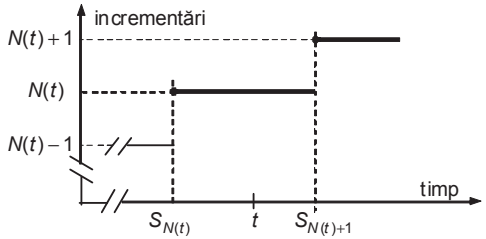


Figura 9.6.3: Momentele ultimei şi primei sosiri faţă de momentul *t*

Rata medie a incrementărilor unui proces de numărare $N(t)$ reprezintă valoarea către care tinde raportul $N(t)/t$, atunci când timpul t devine foarte mare ($t \to \infty$). Pentru determinarea acestei valori se pleacă de la următorul şir de inegalităţi:

$$S_{N(t)} \le t < S_{N(t)+1} \qquad (9.6.11)$$

în care, conform cu figura 9.6.3, $S_{N(t)}$ reprezintă timpul ultimei sosiri înainte de depăşirea timpului t, iar $S_{N(t)+1}$ reprezintă timpul primei sosiri după depăşirea timpului *t*. Împărţind inegalităţile cu $N(t)$ obţinem relaţiile:

$$\frac{S_{N(t)}}{N(t)} \le \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)} \qquad (9.6.12)$$

prin intermediul cărora, aplicând metoda "cleştelui", putem determina limita raportului $t/N(t)$. Astfel, ţinând seama că:

$$S_{N(t)} = \sum_{j=1}^{N(t)} T_j$$

unde T_j sunt variabile aleatorii i.i.d. care reprezintă timpul dintre două sosiri succesive, putem scrie primul termen din expresia (9.6.12) sub forma:

$$\frac{S_{N(t)}}{N(t)} = \frac{1}{N(t)} \sum_{j=1}^{N(t)} T_j$$

care, conform legii tari a numerelor mari, tinde, cu probabilitatea 1, către $\mathrm{E}[T]$, atunci când $N(t) \to \infty$ (prin faptul că $t \to \infty$). În privinţa celui de al treilea termen din (9.6.12), acesta poate fi scris sub forma:

$$\frac{S_{N(t)+1}}{N(t)} = \frac{S_{N(t)+1}}{N(t)+1} \cdot \frac{N(t)+1}{N(t)}$$

ceea ce conduce la aceeaşi concluzie, dat fiind faptul că al doilea factor din membrul drept al egalităţii tinde aproape sigur către valoarea 1 când $t \to \infty$.

Prin urmare, la limită, când $t \to \infty$, raportul $t/N(t)$ se găseşte într-o vecinătate infinitezimală în jurul mărimii $\mathrm{E}[T]$, adică rata medie este dată de următoarea relaţie:

$$\lambda = \lim_{t \to \infty} N(t)/t \to 1/\mathrm{E}[T] \qquad (9.6.13)$$

capitolul 10

Lanţuri Markov

În general, variabilele aleatorii dintr-o familie care defineşte un proces stocastic sunt ***interdependente*** din punct de vedere statistic. Ca urmare, relaţiile de interdependenţă ale variabilelor pot constitui un criteriu pe baza căruia procesele aleatorii pot fi clasificate.

Astfel [Mih78], un proces aleatoriu, $X(t)$, este un ***proces Markov de ordinul 1*** *dacă* **viitorul** *său depinde doar de* **prezent**, *fiind deci independent de trecut* (desigur că există şi procese Markov de ordin superior, care extind relaţia de dependenţă şi în trecutul mai mult sau mai puţin apropiat, dar a căror abordare nu prezintă însă interes pentru obiectul acestei lucrări).

Exprimată sub formă analitică, pentru orice momente de timp $t_1 < t_2 < \cdots < t_k < t_{k+1}$, condiţia precizată anterior devine:

- dacă **procesul este evaluat discret**:

$$P\left[\underbrace{X(t_{k+1}) = x_{k+1}}_{\text{va fi}} \,\middle|\, \underbrace{X(t_k) = x_k}_{\text{este}}, \underbrace{X(t_{k-1}) = x_{k-1},\ \ldots, X(t_1) = x_1}_{\text{a fost}}\right] = P\left[X(t_{k+1}) = x_{k+1} \,\middle|\, X(t_k) = x_k\right] \tag{10.1}$$

- dacă **procesul este evaluat continuu**:

$$P\left[a < X(t_{k+1}) \le b \,\middle|\, X(t_k) = x_k, \cdots, X(t_1) = x_1\right] = P\left[a < X(t_{k+1}) \le b \,\middle|\, X(t_k) = x_k\right] \tag{10.2}$$

Atunci când funcţiile de distribuţie de legătură ale procesului $X(t)$ sunt continue, ultima expresie este echivalentă cu:

$$f_{X(t_{k+1})}\left(x_{k+1} \middle| X(t_k) = x_k, \cdots, X(t_1) = x_1\right) = f_{X(t_{k+1})}\left(x_{k+1} \middle| X(t_k) = x_k\right) \tag{10.3}$$

Valoarea, pe care un proces Markov, $X(t)$, o ia la un moment dat şi faţă de care se definesc probabilităţile sus menţionate, se numeşte ***starea procesului*** la momentul t. Desigur, mulţimea stărilor este spaţiul variabilei aleatorii obţinute prin eşantionarea procesului la momentul t.

Un proces Markov care este evaluat discret, pe o mulţime finită sau infinit numărabilă, se numeşte ***lanţ Markov***. Probabilităţile care caracterizează stările unui lanţ Markov se notează cu:

$$P\left[X(t) = j\right] = p_j(t) \tag{10.4}$$

unde j este una din stările respectivului proces. Dacă un astfel de proces este şi staţionar, atunci ***probabilitatea (absolută) a stării*** este independentă faţă de timp; de aceea pentru aceste procese se poate folosi notaţia simplificată:

$$p_j(t) \overset{\text{notat}}{=} p_j \tag{10.5}$$

Aplicaţia 10.1

Arătaţi (şi motivaţi) care dintre procesele analizate în capitolul anterior sunt procese Markov.

* * *

Aplicaţia 10.2

Exprimaţi funcţia de distribuţie de legătură de ordinul 3 care caracterizează un proces Markov prin intermediul probabilităţilor condiţionate. Generalizaţi rezultatul.

Rezolvare:

$$\begin{aligned} &P\left[X(t_1) = x_1, X(t_2) = x_2, X(t_3) = x_3\right] = \\ &= P\left[X(t_3) = x_3 \middle| X(t_1) = x_1, X(t_2) = x_2\right] \cdot P\left[X(t_1) = x_1, X(t_2) = x_2\right] = \\ &= P\left[X(t_1) = x_1\right] \cdot P\left[X(t_2) = x_2 \middle| X(t_1) = x_1\right] \cdot P\left[X(t_3) = x_3 \middle| X(t_2) = x_2\right] \end{aligned} \tag{10.6}$$

* * *

Relaţia (10.6), precum şi echivalenta sa generalizată, sugerează un fapt care se va confirma în continuare şi anume: un proces aleatoriu de tip Markov poate fi definit complet prin intermediul probabilităţilor stărilor iniţiale şi a probabilităţilor de tranziţie de la o stare la alta între oricare două momente de timp; de exemplu, în cazul relaţiei menţionate:

$$p_{x_1, x_2} = P\left[X(t_2) = x_2 \middle| X(t_1) = x_1\right].$$

10.1 Lanţuri Markov discrete

Probabilitatea de legătură a primelor $n+1$ variabile aleatorii aparţinând unei familii care alcătuieşte un lanţ Markov discret în timp **DTMC** (***D**iscret **T**ime **M**arkov **C**hain*), X_n, ce începe cu $n=0$, are, prin extinderea relaţiei (10.6), următoarea expresie:

$$P\left[X_0=i_0, X_1=i_1,\cdots, X_n=i_n\right]= \\ =P\left[X_0=i_0\right]\cdot P\left[X_1=i_1 \middle| X_0=i_0\right]\cdot\ldots\cdot P\left[X_n=i_n \middle| X_{n-1}=i_{n-1}\right] \quad (10.1.1)$$

în care factorii de genul $P\left[X_k=i_k \middle| X_{k-1}=i_{k-1}\right]$, cu $0<k\le n$, reprezintă ***probabilităţile (condiţionate) de tranziţie*** din starea i_{k-1} în starea i_k, odată cu ***avansarea*** procesului cu un pas (de la momentul $k-1$ la momentul k), iar factorul $P\left[X_0=i_0\right]$ este probabilitatea ca procesul să ***plece*** din starea i_0.

În general, probabilităţile de tranziţie care caracterizează un anumit lanţ Markov pot depinde de momentul de referinţă al avansării. Însă, atunci când probabilităţile de tranziţie nu depind de acest moment (este cazul majorităţii proceselor analizate pe parcursul acestei lucrări), spunem că procesul respectiv este un ***proces omogen în timp***. În această situaţie, relaţia (10.1.1) devine:

$$P\left[X_0=i_0, X_1=i_1,\cdots, X_n=i_n\right]=p_{i_0}(0)\cdot p_{i_0,i_1}\cdot\ldots\cdot p_{i_{n-1},i_n} \quad (10.1.2)$$

unde s-au făcut notaţiile:

$$P\left[X_k=i_k \middle| X_{k-1}=i_{k-1}\right]\overset{\text{notat}}{=}p_{i_{k-1},i_k}(1),\ 0<k\le n \quad (10.1.3)$$

$$P\left[X_0=i_0\right]\overset{\text{notat}}{=}p_{i_0}(0) \quad (10.1.4)$$

În cadrul notaţiei $p_{i_{k-1},i_k}(1)$, cifra dintre paranteze reprezintă numărul de paşi între momentele luate în considerare (anume $k-1$ şi k), iar în cadrul notaţiei $p_{i_0}(0)$ cifra dintre paranteze reprezintă momentul de referinţă. Atunci când se ia în considerare un singur pas, paranteza poate fi omisă, recurgându-se la notaţia simplificată:

$$p_{i_{k-1},i_k}(1)\overset{\text{notat}}{=}p_{i_{k-1},i_k} \quad (10.1.5)$$

deja folosită în relaţia (10.1.2).

Ansamblul probabilităţilor de tranziţie definite prin relaţia (10.1.3) alcătuiesc ***matricea probabilităţilor de tranziţie*** la avansarea cu un pas a unui lanţ Markov discret în timp, ce este organizată şi notată ca:

$$\mathbf{P}(1)=\begin{bmatrix} p_{0,0} & p_{0,1} & p_{0,2} & \cdots \\ p_{1,0} & p_{1,1} & p_{1,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ p_{i,0} & p_{i,1} & p_{i,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (10.1.6)$$

Desigur că elementele fiecărei linii din matricea de tranziţii $\mathbf{P}(1)$ respectă relaţia de normare:

$$\sum_j P\left[X_{n+1}=j \middle| X_n=i\right]=\sum_j p_{i,j}=1 \quad (10.1.7)$$

care este exprimarea matematică a certitudinii că procesul Markov tranzitează şi deci se află oricând într-una din stările sale. În plus, şi în acest caz se poate face o simplificare de notaţie, în sensul că nu se mai precizează numărul paşilor de tranziţie, decât dacă el este supraunitar, altfel $\mathbf{P}(1)\overset{\text{notat}}{=}\mathbf{P}$.

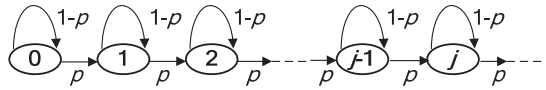


Figura 10.1.1: Diagrama de tranziţii a procesului binomial de numărare

Pentru o prezentare sugestivă a tranziţiilor care caracterizează un lanţ Markov discret, se utilizează reprezentări grafice de genul celei din figura 10.1.1, care corespunde unui proces binomial de numărare (vezi aplicaţia 6.2.2).

Aplicaţia 10.1.1

Scrieţi, relativ la diagrama din figura 10.1.1, matricea probabilităţilor de tranziţii.

* * *

Aplicaţia 10.1.2

În urma observării modului de manifestare a zgomotului într-un canal binar, s-a desprins următoarea regulă de perturbare a simbolurilor recepţionate:

- dacă un bit nu a fost perturbat, următorul bit poate fi perturbat cu probabilitate a,
- iar, dacă a fost perturbat, următorul bit poate fi neperturbat cu probabilitate b.

Reprezentaţi diagrama de tranziţii a acestui proces şi scrieţi matricea de tranziţie corespunzătoare.

Indicaţie: Considerând zgomotul o secvenţă binară care se însumează modulo 2, bit cu bit, cu secvenţa generată la emisie, înseamnă că bitul de informaţie va fi perturbat doar dacă *bitul de zgomot* corespunzător este 1. Deci, procesul de perturbare se desfăşoară conform diagramei din figura 10.1.2.

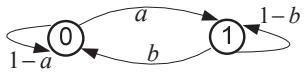


Figura 10.1.2: Diagrama de tranziţii a unui lanţ Markov cu două stări

* * *

Relaţia (10.1.1), pentru probabilităţile de legătură a $n+1$ variabile aleatorii, corespunde unui caz particular ce ia în considerare că cele $n+1$ momente de sondare a unui lanţ Markov discret, X_n, sunt ***succesive***. Însă, cazul general se referă la momente care nu sunt obligatoriu succesive, de exemplu k şi $k+n$, $(n>1)$, fapt pentru care este necesară determinarea probabilităţilor de tranziţie la o avansare a procesului cu mai mulţi paşi; aceste probabilităţi sunt definite conform relaţiei:

$$p_{i,j}(n) = P\left[X_{k+n} = j \,\middle|\, X_k = i\right], \text{ pentru } \forall k, n \geq 2, i \geq 0, j \geq 0 \quad (10.1.8)$$

În dorinţa exprimării acestor probabilităţi în funcţie de probabilităţile de tranziţie la avansarea cu un singur pas, relaţiile căutate se determină aplicând metoda inducţiei, plecând de la stabilirea expresiei pentru probabilităţile de tranziţie de-a lungul a 2 paşi.

Acest demers ia în considerare doar lanţurile Markov care prezintă proprietatea de omogenitate temporală, ce se exprimă analitic prin relaţia următoare:

$$p_{i,j}(n) = P\left[X_{k+n} = j \,\middle|\, X_k = i\right] = P\left[X_n = j \,\middle|\, X_0 = i\right] \quad (10.1.9)$$

cu $n \geq 0$ şi $k \geq 0$. În aceste condiţii, probabilităţile de tranziţie la avansarea cu 2 paşi, plecând din starea i şi ajungând starea j, *via* starea intermediară k, este dată de relaţia:

$$P\left[X_1 = k, X_2 = j \,\middle|\, X_0 = i\right] = \frac{P\left[X_0 = i, X_1 = k, X_2 = j\right]}{P\left[X_0 = i\right]} = \ldots$$
$$= P\left[X_1 = k \,\middle|\, X_0 = i\right] \cdot P\left[X_2 = j \,\middle|\, X_1 = k\right] = p_{i,k}(1) \cdot p_{k,j}(1) \quad (10.1.10)$$

În general, starea k, prin care procesul *tranzitează* de la starea i la

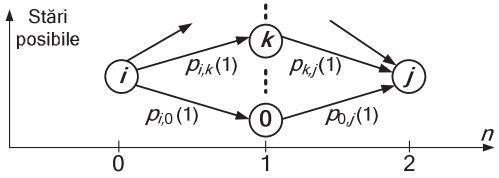


Figura 10.1.3: Graful tranziţiilor prin 2 paşi

starea j, poate fi oarecare (figura 10.1.3).

În consecinţă, probabilitatea de tranziţie în 2 paşi, din starea i în starea j, indiferent prin ce stare intermediară, este dată de suma probabilităţilor tuturor traseelor posibile, adică:

$$p_{i,j}(2) = \sum_k p_{i,k}(1) \cdot p_{k,j}(1), \ \forall i,j \quad (10.1.11)$$

ceea ce matriceal se exprimă prin: $\mathbf{P}(2) = \mathbf{P}(1) \cdot \mathbf{P}(1) = \mathbf{P}^2$

Aplicând metoda inducţiei complete, exprimată prin intermediul relaţiei $\mathbf{P}(n) = \mathbf{P}(n-1) \cdot \mathbf{P}$, se obţine expresia:

$$\mathbf{P}(n) = \mathbf{P}^n \quad (10.1.12)$$

care evidenţiază faptul că matricea probabilităţilor de tranziţie de-a lungul unui număr oarecare de paşi se poate determina în funcţie de matricea probabilităţilor de tranziţie după un pas, care astfel, împreună cu probabilităţile iniţiale de stare, sunt suficiente pentru a specifica complet un lanţ Markov discret.

Aplicaţia 10.1.3

Determinaţi matricea de tranziţie la avansarea cu doi paşi a procesului analizat în cadrul aplicaţiei 10.1.1.

* * *

Un lanţ Markov discret poate intra la un anumit pas, n, într-una din stările sale posibile conform unei distribuţii probabilistice exprimată prin intermediul unui vector, numit ***vectorul probabilităţilor de stare*** la momentul n:

$$\mathbf{p}(n) = \left\{p_j(n)\right\} \quad (10.1.13)$$

Elementele $p_j(n)$ ale vectorului $\mathbf{p}(n)$ sunt probabilităţile ca procesul să se găsească în starea j la pasul n şi se leagă de elementele vectorului $\mathbf{p}(n-1)$, corespunzător pasului precedent $n-1$, prin relaţia:

$$p_j(n) = \sum_i P\left[X_n = j \,\middle|\, X_{n-1} = i\right] \cdot P\left[X_{n-1} = i\right] = \sum_i p_{i,j} \cdot p_i(n-1) \quad (10.1.14a)$$

care, tradusă matriceal, devine relaţia de recurenţă:

$$\mathbf{p}(n) = \mathbf{p}(n-1) \cdot \mathbf{P} \quad (10.1.14b)$$

prin intermediul căreia se determină relaţia care leagă vectorul $\mathbf{p}(n)$ de vectorul probabilităţilor iniţiale de stare, $\mathbf{p}(0)$:

$$\mathbf{p}(n) = \mathbf{p}(0) \cdot \mathbf{P}^n \quad (10.1.15)$$

Relaţia (10.1.15) evidenţiază un lucru deja menţionat, dar care merită subliniat, şi anume că matricea $\mathbf{P}$, împreună cu vectorul $\mathbf{p}(0)$ al probabilităţilor iniţiale de stare definesc complet un lanţ Markov evaluat discret.

Aplicaţia 10.1.4

Considerând procesul din aplicaţia 10.1.2, în care $a=1/10$ şi $b=1/5$, să se determine: - matricea de tranziţie $\mathbf{P}(n)$ pentru $n=2,4,8,16,\cdots,\infty$;

- probabilităţile de stare pentru $n\to\infty$ ştiind că $\mathbf{p}(0)=(1/5;4/5)$.

Rezolvare: Ştiind că, în cazul de faţă: $P=\begin{pmatrix}1-a & a\\ b & 1-b\end{pmatrix}$ şi aplicând cunoştinţele din algebră [Mih78] se poate demonstra că:

$$\mathbf{P}(n)=\mathbf{P}^{\,n}=\frac{1}{a+b}\cdot\begin{pmatrix}b & a\\ b & a\end{pmatrix}+\frac{(1-a-b)^n}{a+b}\cdot\begin{pmatrix}a & -a\\ -b & b\end{pmatrix}$$

ceea ce, la limită, conduce la rezultatul: $\lim\limits_{n\to\infty}\mathbf{P}(n)=\frac{1}{a+b}\cdot\begin{pmatrix}b & a\\ b & a\end{pmatrix}$.

Ducând la limită relaţia (10.1.15), se obţine că:

$$\lim_{n\to\infty}\mathbf{p}(n)=\left(p_0(0),p_1(0)\right)\cdot\lim_{n\to\infty}\mathbf{P}(n)=\left(\frac{b}{a+b},\frac{a}{a+b}\right)$$

* * *

Aplicaţia anterioară pune în evidenţă faptul că unele lanţuri Markov intră într-un regim staţionar de manifestare după un număr mare de paşi. În acest caz, când $n\to\infty$, coloanele matricei de tranzit de-a lungul a n paşi tind să conţină aceleaşi valori, adică:

$$\lim_{n\to\infty}p_{i,\,j}(n)=p_j\ ,\ \forall\, i \qquad (10.1.16)$$

ceea ce, textual, s-ar traduce prin aceea că indiferent de stare i în care s-ar ***găsi*** la un moment dat, astfel de procese ***ajung*** cu aceeaşi probabilitate într-o stare j după un număr mare de paşi intermediari. Prin urmare, pentru astfel de procese ecuaţia (10.1.15) devine, la limită:

$$\lim_{n\to\infty}p_j(n)=\sum_i\ p_j\cdot p_i(0)=p_j\text{, pentru }\forall\, j \qquad (10.1.17)$$

care, în consecinţă, arată că după un număr mare de paşi, se ajunge la un echilibru statistic, în care ***probabilităţile de stare rămân constante, valorile respective fiind independente de probabilităţile iniţiale de stare***.

În cazul lanţurilor Markov discrete care intră în echilibru statistic, noţiunea de *probabilităţi de stare* este completată cu atributul *staţionar* rezultând conceptul de ***probabilităţi staţionare de stare***. Expresia acestor probabilităţi, p_j, se determină prin intermediul ecuaţiei matriceale care descrie tranziţia cu un pas de la un moment *n* îndepărtat de origine, adică:

$$\mathbf{p}=\mathbf{p}\cdot\mathbf{P}\text{, unde }\mathbf{p}=\left\{p_j\right\} \qquad (10.1.18)$$

Acest sistem de ecuaţii este *liniar dependent*, motiv pentru care rezolvarea sa impune înlocuirea uneia dintre ecuaţiile sale cu relaţia de normare:

$$\sum_j\ p_j=1 \qquad (10.1.19)$$

Aplicaţia 10.1.5

Determinaţi probabilităţile staţionare de stare corespunzătoare procesului din aplicaţia 10.1.2.

* * *

Aplicaţia 10.1.6

Arătaţi că un lanţ Markov, care după un anumit număr de paşi ($n\to\infty$) intră, statistic, într-un regim staţionar, este un proces asimptotic staţionar.

Indicaţie: se pleacă de la expresiile probabilităţii stării la momentul n_0+m şi ale probabilităţilor de tranziţie şi se verifică dacă expresia de calcul a probabilităţii de legătură:

$$P\left[X_{n_0+m}=i_0,\,X_{n_1+m}=i_1,\cdots,\,X_{n_k+m}=i_k\right],\ m>0\ \text{ şi }\ \forall k$$

nu depinde de valoarea deplasării *m*. Secvenţa precizată $n_0<n_1<\cdots<n_k$ se află în domeniul de staţionaritate.

* * *

Lanţurile Markov care tind către echilibru statistic devin ***staţionare în sens strict*** dacă, chiar din primul pas, $n=0$, probabilităţile de stare sunt egale cu probabilităţile staţionare de stare, $\mathbf{p}(0)=\mathbf{p}$. Într-adevăr, în acest caz:

$$\mathbf{p}(n)=\mathbf{p}\cdot\mathbf{P}^{\,n}=\mathbf{p}\ ,\ \forall n \qquad (10.1.20)$$

ceea ce conduce, ţinând cont şi de rezultatul obţinut în aplicaţia precedentă, la relaţia:

$$P\left[X_{n_0+m}=i_0,\,X_{n_1+m}=i_1,\cdots,\,X_{n_k+m}=i_k\right]\overset{\substack{\text{omogenitatea}\\ \text{temporală}\\ (7.1.20)}}{=}\pi_{i_0}\cdot p_{i_0,i_1}(n_1-n_0)\cdot p_{i_1,i_2}(n_2-n_1)\cdot\ldots\cdot p_{i_{k-1},i_k}(n_k-n_{k-1}) \qquad (10.1.21)$$

care arată că probabilităţile de legătură corespunzătoare sunt independente de originea (n_0+m pentru $\forall n_0$ şi $\forall m$) la care se raportează, respectând astfel definiţia proceselor strict staţionare.

În general, nu toate lanţurile Markov discrete în timp *ating* un echilibru statistic, fapt pentru care atât relaţia (10.1.17), cât şi rezultatele derivate din aceasta îşi pierd valabilitatea. În sprijinul acestei observaţii se propune rezolvarea aplicaţiei 10.3.1.

Aplicaţia 10.1.7

Să se realizeze un program în MatLab care să simuleze un lanţ Markov discret cu două stări, în varianta:

i) sub controlul ceasului; ii) sub controlul evenimentelor.

Rezolvare: Diagrama avută în vedere este cea prezentată în figura 10.1.4. Pe

lângă cele două **stări recurente** (în care procesul revine de un număr **infinit** de ori în evoluţia sa), diagrama conţine şi starea de plecare (START), care este o **stare tranzitorie** (în care procesul revine de un număr finit de ori în evoluţia sa, în cazul de faţă: 0).

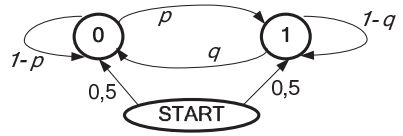


Figura 10.1.4: Lanţ Markov discret cu două stări recurente şi o stare tranzitorie

i) Principial, programul constă într-o bucla care este reluată la intervale precise de timp până la expirarea duratei considerate, de observare a fenomenului.

Concret, schiţa programului cerut arată astfel:

```
 %initializari
P=...;
q=...;
end_sim_time=...;
  % definirea vectorului traiectorie
trace=zeros[2,end_sim_time];
  %fixarea starii initiale
state=0;
if rand()<0.5
 state=1
  %inregistrarea traiectoriei
for clock=1:end_sim_time
 trace[1,clock]=clock;
 trace[2,clock]=state;
 r=rand();
 if state==0
      if r<p
           state=1
 elseif r<q
      state=0
 end %if
end %for
```

ii) Principial, programul constă într-o bucla care este reluată atunci când procesul îşi modifică starea (apare un eveniment), până la expirarea duratei considerate, de observare a fenomenului. Concret, schiţa programului cerut arată astfel:

```
 %initializari
P=...;
q=...;
```



```
end_sim_time=...;
  %definirea vectorului traiectorie
trace=zeros[2,end_sim_time];
  %fixarea starii initiale
state=0;
if rand()<0.5
  state=1
  %se stabileste cand are loc urmatoarea schimbare de
  %stare
if state==0
  time=geomdist(p);
else
  time=geomdist(q)
    %inregistrarea traiectoriei
while(time<end_sim_time)
  if state==0
       state=1;
       trace=[trace; time state];
         %se stabileste momentul urmatoarei tranzitii
       time=time+geomdist(q);
  else
       state=0;
       trace=[trace; time state];
         %se stabileste momentul urmatoarei tranzitii
       time=time+geomdist(p);
  end  %if
end  %while
```

Aplicaţia 10.1.8

Stabiliţi valoarea parametrului `end_sim_time` funcţie de mărimile p şi q, astfel încât, pe durata execuţiei celui de-al doilea program din aplicaţia 10.1.7, să apară, în medie, 10000 de evenimente (tranziţii dintr-o stare în alta).

Indicaţie: se pleacă de la duratele medii de staţionare în fiecare stare, pe baza cărora se stabileşte durata medie a unui ciclu de revenire în aceeaşi stare.

* * *

Aplicaţia 10.1.9

Fixând timpul de simulare conform formulei de calcul stabilite odată cu rezolvarea aplicaţiei 10.1.8, comparaţi cele două programe de simulare din aplicaţia 10.1.7 din punct de vedere al timpului de execuţie, luând în considerare, pentru mărimile p şi q, următoarele perechi de valori: $(1/2, 1/2)$ şi $(1/10, 1/10)$. Justificaţi diferenţele ce apar în ambele cazuri.

* * *

Aplicaţia 10.1.10

Pentru lanţul Markov din figura 10.1.5: i) precizaţi care stări sunt

recurente şi care sunt tranzitorii; ii) determinaţi probabilităţile staţionare de stare dacă: a) $p_{-1}(0)=1$; b) $p_0(0)=1$ sau c) $p_2(0)=1$.

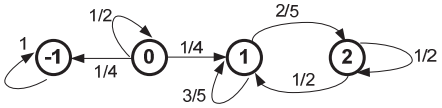


Figura 10.1.5: Lanţ Markov asociat aplicaţiei 10.1.10

Indicaţie: definiţiile privind cele două tipuri de stări sunt date în aplicaţia 10.1.7.

10.2 Lanţuri Markov continue

Probabilităţile de legătură care specifică un lanţ Markov continuu în timp **CTMC** (***C**ontinuous **T**ime **M**arkov **C**hain*) pot fi determinate prin intermediul probabilităţilor condiţionate, de tranziţie între stări, conform relaţiei următoare, obţinută prin generalizarea relaţiei (10.6):

$$\begin{aligned} &P\left[X(t_1)=i_1, X(t_2)=i_2, \cdots, X(t_{k+1})=i_{k+1}\right] \\ &= P\left[X(t_1)=i_1\right]\cdot P\left[X(t_2)=i_2 \middle| X(t_1)=i_1\right]\cdot\cdots\cdot P\left[X(t_{k+1})=i_{k+1} \middle| X(t_k)=i_k\right] \end{aligned} \quad (10.2.1)$$

unde $t_1, t_2, \cdots, t_{k+1}$ sunt $k+1$ momente arbitrare ordonate crescător în timp, iar $P\left[X(t_1)=i_1\right]$ este probabilitatea ca la momentul t_1 lanţul să se găsească în starea i_1.

Deci, şi în această situaţie, ceea ce ne interesează este stabilirea unor relaţii de calcul pentru probabilităţile de tranziţie care, de data aceasta, iau în considerare momente arbitrare de timp, de exemplu s şi $s+t$, ce aparţin unui domeniu continuu de valori.

Concentrându-ne atenţia, şi în acest caz, asupra lanţurilor care respectă proprietatea de omogenitate temporală, adică:

$$p_{i,j}(t) \overset{notat}{=} P\left[X(s+t)=j \middle| X(s)=i\right] = P\left[X(t)=j \middle| X(0)=i\right],\ t\geq 0, \forall s \quad (10.2.2)$$

putem defini matricea corespunzătoare a probabilităţilor de tranziţie de-a lungul unui interval t:

$$\mathbf{P}(t)=\left\{p_{i,j}(t)\right\} \quad (10.2.3)$$

De remarcat că, pentru $t=0$, avem $p_{i,i}(0)=1$ şi $p_{i,j}(0)=0$ pentru orice $i\neq j$, adică matricea probabilităţilor de tranziţie capătă forma unei matrice identică (matricea unitate):

$$\mathbf{P}(0)=\mathbf{I} \quad (10.2.4)$$

Aplicaţia 10.2.1

Scrieţi matricea probabilităţilor de tranziţie de-a lungul unui interval, t, în cazul procesului Poisson de numărare, de rată medie λ. Care este conţinutul acestei matrice în cazul în care se consideră un interval δ suficient de mic pentru a se respecta inegalitatea $\lambda\delta<<1$.

Indicaţie: Proprietatea de staţionaritate a incrementării, proprie proceselor Poisson de numărare, este echivalentă cu cea a omogenităţii temporale ceea ce permite scrierea egalităţii: $p_{i,j}(t) \overset{\text{omogenitatea temporală}}{=} p_{0,j-i}(t)=\frac{(\lambda\cdot t)^{j-i}}{(j-i)!}e^{-\lambda\cdot t}$, pentru orice $j\geq i$, care reprezintă de fapt probabilitatea ca într-un interval t să apară $j-i$ evenimente. Prin urmare, matricea căutată are expresia:

$$\mathbf{P}(t)=\begin{bmatrix} 1 & \lambda\cdot t & (\lambda\cdot t)^2/2! & \cdots & \cdots \\ 0 & 1 & \lambda\cdot t & (\lambda\cdot t)^2/2! & \cdots \\ 0 & 0 & 1 & \lambda\cdot t & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}\cdot\exp(-\lambda\cdot t)$$

Atunci când $\lambda t<<1$, se poate utiliza aproximaţia $\exp(-\lambda\cdot t)\cong 1-\lambda\cdot t$, motiv pentru care, pentru intervale δ suficient de mici faţă de raportul $1/\lambda$, matricea anterioară devine:

$$\mathbf{P}(\delta)\cong\begin{bmatrix} 1-\lambda\cdot\delta & \lambda\cdot\delta & 0 & 0 & \cdots \\ 0 & 1-\lambda\cdot\delta & \lambda\cdot\delta & 0 & \cdots \\ 0 & 0 & 1-\lambda\cdot\delta & \lambda\cdot\delta & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Se pune astfel în evidenţă faptul că, pe intervale foarte scurte, incrementările cu mai mult de o unitate ale procesului Poisson de numărare sunt foarte puţin probabile, adică au o probabilitate practic nulă. Anterior această afirmaţie a fost luată ca ipoteză, atunci când s-a prezentat procesul Poisson ca exemplu de proces aleatoriu continuu.

* * *

Lipsa dependenţei de trecut a proceselor Markov (de ordinul întâi), la care contează doar prezentul, face ca probabilitatea condiţionată de staţionare a unui astfel de proces încă t unităţi temporale (u.t.) în aceeaşi stare i, în care se găseşte deja de s u.t. să nu depindă de valoarea *s* (considerată ca "moment prezent" raportat la momentul intrării în starea respectivă), adică:

$$P\left[T_i>t+s \middle| T_i>s\right]=P\left[T_i>t \middle| T_i>0\right]=P\left[T_i>t\right]$$

unde T_i este variabila aleatorie pozitivă ce reprezintă ***timpul de staţionare*** în starea i.

Această proprietate poartă denumirea de ***lipsă a memoriei*** şi este respectată doar de o singură variabilă aleatorie continuă şi anume de către variabila exponenţială. În consecinţă timpul de staţionare urmează o distribuţie exponenţială, având expresia:

$$P[T_i > t] = \exp(-\mu_i \cdot t) \qquad (10.2.5)$$

în care $1/\mu_i$ reprezintă media acestui timp, care poate avea valori diferite pentru diferitele stări ale procesului.

Proprietatea demonstrată mai sus permite interpretarea unui lanţ Markov continuu în timp ca o interpolare de ordinul zero (figura 10.2.1) a unui lanţ discret, [Doy89], numit ***lanţ Markov asociat*** (***încuibărit***, în limba engleză ***"embadded"***) în care tranziţiile $i \to j$ se desfăşoară cu probabilităţi $q_{i,j}$, iar intervalul T_j dintre două momente de avans ale lanţului discret urmează o distribuţie exponenţială de parametru μ_j şi reprezintă durata de staţionare a procesului continuu în starea j.

În acest context, probabilitatea ca un lanţ Markov continuu să rămână într-o stare oarecare, i, cel puţin pe durata unui interval infinitezimal, δ, este:

$$P[T_i > \delta] = e^{-\mu_i \cdot \delta} = 1 - \frac{\mu_i \cdot \delta}{1!} + \frac{(\mu_i \cdot \delta)^2}{2!} - \frac{(\mu_i \cdot \delta)^3}{3!} + \ldots = 1 - \mu_i \cdot \delta + o(\delta) \qquad (10.2.6)$$

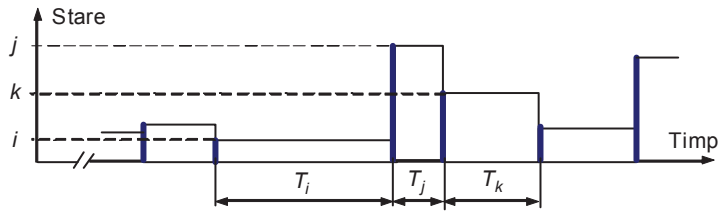


Figura 10.2.1: Legătura dintre un lanţ Markov continuu şi unul asociat - exemplificare

În general, prin $o(x)$ se aproximează orice funcţie $g(x)$, când $x \to 0$, pentru care este adevărat că $\lim_{x \to 0} g(x)/x = 0$.

În principiu, procesul poate să şi părăsească starea *i* înaintea expirării intervalului δ, acest eveniment având loc cu probabilitatea $P[T_i \le \delta] \cong \mu_i \cdot \delta$.

Deci probabilitatea de tranziţie, $p_{i,i}(\delta)$, implică atât staţionarea continuă în respectiva stare pe întreaga durată δ, cât şi evenimente de genul $\{$*pe durata* δ*, procesul tranzitează prin mai multe stări, plecând şi ajungând în aceeaşi stare*$\}$. Însă astfel de evenimente au probabilităţi de apariţie neglijabile.

În sprijinul acestei afirmaţii, poate fi exemplificat cazul tranziţiei pe *traseul* $i \to j \to i$, pe durata δ, a cărei probabilitate, $P[T_i + T_j + T_i' < \delta]$, unde T_i, T_j şi T_i' sunt timpii de staţionare în cele trei stări (considerate, pentru simplitatea justificării, de medii egale, μ), este dată de distribuţia *m*-Erlang, cu $m = 3$:

$$P[T_i + T_j + T_i' < \delta] = 1 - \sum_{k=0}^{2} \frac{(\mu \cdot \delta)^k}{k!} \cdot e^{-\mu \cdot \delta} = o(\delta)$$

În consecinţă, probabilitatea staţionării în starea i pe întreaga durată δ, $p_{i,i}(\delta)$, are expresia:

$$p_{i,i}(\delta) = 1 - \mu_i \cdot \delta \qquad (10.2.7)$$

din care se poate deduce că probabilitatea ca procesul $X(t)$ să părăsească starea *i* pe parcursul respectivului interval este:

$$1 - p_{i,i}(\delta) = \mu_i \cdot \delta \qquad (10.2.8)$$

Semnificaţia relaţiei (10.2.8) şi expresia termenului drept al acesteia dau parametrului μ_i sensul de ***rată de părăsire a stării*** i de către procesul în cauză.

Aşa cum s-a menţionat anterior, odată părăsită o stare *i*, procesul intră cu probabilitatea $q_{i,j}$ într-o altă stare j, fapt pentru care:

$$p_{i,j}(\delta) = [1 - p_{i,i}(\delta)] \cdot q_{i,j} = \mu_i q_{i,j} \cdot \delta = \gamma_{i,j} \cdot \delta \qquad (10.2.9)$$

în care, prin analogie, mărimea $\gamma_{i,j}$ reprezintă, pentru procesul în cauză $X(t)$, ***rata de tranziţie*** *din starea* i *în starea* j.

Cu ajutorul probabilităţilor de tranziţie, se pot determina probabilităţile de stare, $P[X(t) = j] \overset{notat}{=} p_j(t)$, plecând de la următorul şir de egalităţi (valabil, desigur, în cazul unor intervale δ foarte mici):

$$\begin{aligned} p_j(t+\delta) &= P[X(t+\delta) = j] \\ &= \sum_i P[X(t+\delta) = j \,|\, X(t) = i] \cdot P[X(t) = i] \\ &= \sum_i p_{i,j}(\delta) \cdot p_i(t) \end{aligned} \qquad (10.2.10)$$

în care, expresia din ultima egalitate corespunde diagramei de tranziţii din figura 10.2.2.

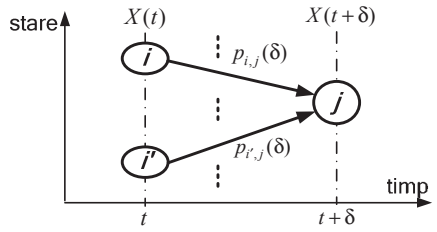


Figura 10.2.2: Exemple de tranziţii spre starea *j*

Scăzând în relaţia (10.2.10) termenul $p_j(t)$ şi împărţind ambii termeni ai ultimei egalităţi cu δ se obţine expresia:

$$\frac{p_j(t+\delta)-p_j(t)}{\delta}=\sum_{i\neq j}\frac{p_{i,j}(t)}{\delta}\cdot p_i(t)+\frac{p_{i,j}(\delta)-1}{\delta}\cdot p_j(t) \quad (10.2.11)$$

care la limită, când $\delta\to 0$, şi făcând substituţia $-\mu_i \overset{\text{notat}}{=} \gamma_{i,i}$ în relaţia (10.2.8), capătă forma cunoscută a setului de ***ecuaţii Chapman-Kolmogorov*** (***ecuaţiile viitorului***) ce sunt caracteristice lanţurilor Markov continue în timp:

$$p'_j(t)=\sum_i \gamma_{i,j}\cdot p_i(t) \quad (10.2.12)$$

Sub formă matriceală, expresia (10.2.12) devine:

$$\mathrm{d}\,\mathbf{p}(t)/\mathrm{d}t=\mathbf{p}(t)\cdot\mathbf{Q} \quad (10.2.13)$$

unde: $\mathbf{p}(t)=\{p_0(t),\ p_1(t),\cdots\}$ este vectorul probabilităţilor de stare la momentul t, iar $\mathbf{Q}=\{\gamma_{i,j}\}$ este ***matricea (generatoare) infinitezimală***.

Determinarea efectivă a probabilităţilor de stare constă în rezolvarea sistemului de ecuaţii diferenţiale (10.2.13), pentru care se stabilesc anumite condiţii iniţiale $\mathbf{p}(0)=\{p_0(0),\ p_1(0),\cdots\}$ şi a cărui soluţie are forma generală:

$$\mathbf{p}(t)=\mathbf{p}(0)\cdot\exp(\mathbf{Q}\cdot t) \quad (10.2.14)$$

În funcţie de alegerea probabilităţilor iniţiale de stare (anume a elementelor vectorului $\mathbf{p}(0)$ de stare), se poate determina şi matricea de tranziţie $\mathbf{P}(t)$. Astfel, dacă considerăm că procesul *porneşte* din starea i, adică $p_i(0)=1$ şi $p_j(0)=0$, $\forall\, j\neq i$, atunci soluţia sistemului de ecuaţii diferenţiale reprezintă componentele liniei i din matricea $\mathbf{P}(t)$, adică probabilităţile $p_{i,j}(t)$ ale tuturor tranziţiilor $i\to j$.

Aplicaţia 10.2.2

Linia unui abonat telefonic se poate găsi în două stări: *liber* sau *ocupat* (notate 0, respectiv 1). Ştiind că duratele celor două tipuri de stări urmează distribuţii exponenţiale de medie $1/\alpha=5/4$, respectiv $1/\beta=5$, să se determine probabilităţile de stare: $p_0(t)$ şi $p_1(t)$ în funcţie de probabilităţile iniţiale de stare $p_0(0)=0,3$ şi $p_1(0)=0,7$.

Rezolvare: Expresiile de calcul al ratelor de tranziţie sunt:

$$\gamma_{0,1}=\mu_0\cdot q_{0,1}=\alpha;\quad \gamma_{0,0}=-\mu_0=-\alpha;$$
$$\gamma_{1,0}=\mu_1\cdot q_{1,0}=\beta;\quad \gamma_{1,1}=-\mu_1=-\beta.$$

Introducându-le în sistemul de ecuaţii diferenţiale (ale viitorului) se obţine:

$$\begin{cases} p'_0(t)=-\alpha\cdot p_0(t)+\beta\cdot p_1(t)\\ p'_1(t)=\alpha\cdot p_0(t)-\beta\cdot p_1(t)\end{cases}$$

la care se adaugă relaţia de normare: $p_0(t)+p_1(t)=1$.

Făcând substituţiile de rigoare, se obţine următoarea ecuaţie diferenţială de ordinul întâi:

$$p'_0(t)+(\alpha+\beta)\cdot p_0(t)=\beta,$$

a cărei soluţie generală este:

$$p_0(t)=\beta/(\alpha+\beta)+C\cdot\mathrm{e}^{-(\alpha+\beta)\cdot t}.$$

Constanta C se determină particularizând soluţia pentru momentul $t=0$, obţinându-se în final soluţiile:

$$p_0(t)=\frac{\beta}{\alpha+\beta}+\left[p_0(0)-\frac{\beta}{\alpha+\beta}\right]\cdot\mathrm{e}^{-(\alpha+\beta)\cdot t}=0,2+0,1\cdot\mathrm{e}^{-t}$$
$$p_1(t)=0,8-0,1\cdot\mathrm{e}^{-t}$$

* * *

Unele lanţuri Markov continue pot atinge, după un interval suficient de mare de timp ($t\to\infty$), un **echilibru statistic** în care probabilităţile de stare se *stabilizează* la anumite valori, care nu depind de valorile probabilităţilor iniţiale de stare (vezi cazul aplicaţiei anterioare în care, de exemplu, $\lim_{t\to\infty} p_0(t)=\beta/(\alpha+\beta)$).

Deci, dacă $p_j(t)\to p_j$, atunci înseamnă că $p'_j(t)\to 0$, ceea ce conduce la următorul sistem de ecuaţii, derivat din relaţia (10.2.12):

$$\sum_i \gamma_{i,j}\cdot p_i=0,\ \forall j \quad (10.2.15)$$

sau, la nivel matriceal:

$$\mathbf{0}=\mathbf{p}\cdot\mathbf{Q} \quad (10.2.16)$$

cu $\mathbf{p}=\{p_i\}$ reprezentând **vectorul probabilităţilor staţionare de stare**.

Ţinând cont că $\gamma_{j,j}=-\mu_j$, relaţia anterioară devine:

$$\mu_j \cdot p_j = \sum_{i \neq j} \gamma_{i,j} \cdot p_i \,, \ \forall j \qquad (10.2.17)$$

şi cum $\mu_j = \sum_{i \neq j} \gamma_{j,i}$, ca rată totală de ieşire din starea j, se obţine în final că:

$$\left(\sum_{i \neq j} \gamma_{j,i}\right) \cdot p_j = \sum_{i \neq j} \left(\gamma_{i,j} \cdot p_i\right), \ \forall j \qquad (10.2.18)$$

Sistemul liniar de ecuaţii din relaţia (10.2.18) reprezintă ***ecuaţiile de echilibru global***, fiind exprimarea matematică a faptului că *fluxul* de probabilitate care iese din starea j (adică $\mu_j \cdot p_j$) este egal cu fluxul de probabilitate *generat* de celelalte stări şi având ca destinaţie starea j (de tipul $\gamma_{i,j} \cdot p_i$).

Toate aceste elemente de flux de probabilitate de intrare şi de ieşire dintr-o stare j se regăsesc în graful de tranziţii din figura 10.2.3.

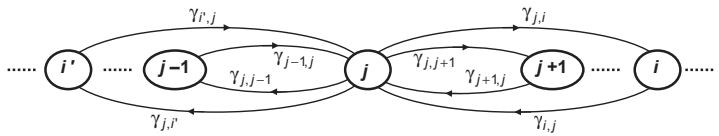


Figura 10.2.3: Echilibrul global al fluxurilor de probabilitate

Întrucât sistemul ecuaţiilor de echilibru global este liniar dependent, determinarea probabilităţii staţionare de stare, $\mathbf{p} = \{p_j\}$, presupune înlocuirea uneia din ecuaţii cu relaţia de normare, $\sum_i p_i = 1$. Chiar şi în urma acestei modificări este posibilă obţinerea unei soluţii nule, $p_j = 0$, $\forall j$, care nu contrazice cu nimic condiţia $p_j'(t) \to 0$, dar care nu este compatibilă cu noţiunea de echilibru statistic. Prin urmare, doar atunci când soluţia diferă de vectorul nul, valorile obţinute reprezintă probabilităţi staţionare de stare.

Ca şi în cazul discret, dacă un lanţ Markov continuu este asimptotic staţionar şi porneşte cu probabilităţile iniţiale de stare egale cu cele staţionare, atunci el devine şi staţionar în sens strict, întrucât $p_j(t) = p_j, \forall t$. Demonstraţia este asemănătoare, bazându-se pe proprietatea de omogenitate temporală şi pe faptul că probabilitatea stării iniţiale, $P[X(t_1) = i_0]$, este aceeaşi pentru orice moment, t_1, ales.

Aplicaţia 10.2.3

Realizaţi un program MatLab care să simuleze un lanţ Markov continuu cu 2 stări.

Indicaţie: se va opta (de ce?) pentru principiul simulării sub controlul evenimentelor, programul fiind asemănător cu cel din aplicaţia 10.1.7.

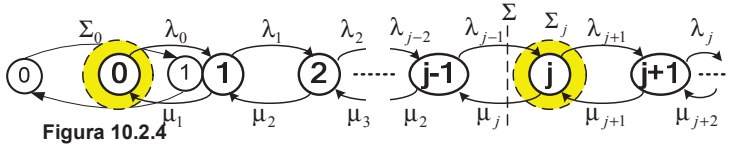


Figura 10.2.4

Figura 10.4.1: Proces de naştere şi moarte: diagrama generală a ratelor de tranziţie

Aplicaţia 10.2.4

Reprezentaţi grafic echilibrul global al fluxurilor probabilistice din *aplicaţia 10.2.2* şi determinaţi probabilităţile staţionare de stare prin intermediul ecuaţiilor de echilibru global.

Indicaţie: Diagrama tranziţiilor corespunzătoare unui lanţ Markov continuu cu 2 stări este prezentată în figura 10.2.4.

10.4 Procese de naştere şi moarte

Un caz particular, de lanţuri Markov continue în timp, utilizat frecvent în modelarea matematică a diverse situaţii reale, îl constituie clasa ***proceselor de naştere şi moarte*** (creştere – descreştere), la care tranziţiile *leagă* numai stări adiacente. În consecinţă, astfel de procese sunt descrise, la modul general, prin intermediul diagramei ratelor de tranziţie din figura 10.4.1, în care λ_i şi μ_i sunt ratele de ieşire din starea i, cu $i \geq 0$.

Diagrama din figura menţionată permite scrierea următoarelor ecuaţii de echilibru global, a căror formă se diferenţiază funcţie de poziţia în diagramă a *conturului* prin care *se schimbă* fluxurile probabilistice:

- suprafaţa Σ_0, ($j = 0$): $\lambda_0 \cdot p_0 = \mu_1 \cdot p_1$ (10.4.1)

- suprafaţa Σ_j, ($j > 0$): $(\lambda_j + \mu_j) \cdot p_j = \lambda_{j-1} \cdot p_{j-1} + \mu_{j+1} \cdot p_{j+1}$ (10.4.2)

Rezolvarea acestui sistem se poate face transcriind relaţiile de mai sus în forma:

$$\lambda_0 \cdot p_0 - \mu_1 \cdot p_1 = 0 \qquad (10.4.3)$$

$$\lambda_{j-1} \cdot p_{j-1} - \mu_j \cdot p_j = \lambda_j \cdot p_j - \mu_{j+1} \cdot p_{j+1} \qquad (10.4.4)$$

exprimare ce pune în evidenţă faptul că, pentru orice $j \geq 0$, toţi termenii $\lambda_j \cdot p_j - \mu_{j+1} \cdot p_{j+1}$ au aceeaşi valoare, stabilită de relaţia (10.4.3), şi anume **zero**. Prin urmare, se poate scrie următoarea relaţie:

$$\underbrace{\lambda_{j-1} \cdot p_{j-1}}_{\text{flux ascendent}} = \underbrace{p_j \cdot \mu_j}_{\text{flux descendent}} \qquad (10.4.5)$$

care precizează faptul că, pentru stabilirea echilibrului este ***necesar***, dar nu

neapărat ***suficient*** (din motive asemănătoare celei expuse anterior, când s-a luat în discuţie soluţia ecuaţiilor de echilibru global), ca fluxurile de probabilitate, *ascendent* şi *descendent*, ce străbat frontiera Σ să fie egale (figura 10.4.1).

Pe baza ultimei relaţii, denumită şi ***ecuaţie de echilibru local***, se poate obţine exprimarea probabilităţii p_j a oricărei stări a procesului în funcţie de probabilitatea p_0 a stării vide, astfel:

$$p_j = \underbrace{r_j \cdot r_{j-1} \cdot \ldots \cdot r_0}_{\overset{notat}{=} R_j} \cdot p_0 = p_0 \cdot \prod_{k=1}^{j} \frac{\lambda_{k-1}}{\mu_k} \qquad (10.4.6)$$

unde $r_k \overset{notat}{=} \lambda_{k-1}/\mu_k$. Conform expresiei (10.4.6), pentru determinarea probabilităţii p_j este suficient să stabilim relaţia de calcul a probabilităţii p_0. Acest lucru se poate face apelând la relaţia de normare, obţinându-se în acest mod ecuaţia:

$$p_0 \cdot \sum_{i=0}^{\infty} R_i = 1, \text{ în care } R_0 = 1. \qquad (10.4.7)$$

Fiind vorba de o sumă infinită, existenţa unei soluţii nenule pentru ecuaţia (10.4.7) impune ca $\sum_{i=0}^{\infty} R_i$ să fie *serie convergentă*. În această situaţie, procesul de naştere şi moarte are probabilităţi staţionare de stare date de relaţia:

$$p_j = \frac{R_j}{\sum_{i=0}^{\infty} R_i} \qquad (10.4.8)$$

În general, atunci când nu se găsesc modalităţi mai simple (precum stabilirea unei formule de recurenţă), rezolvarea ecuaţiilor de echilibru global care caracterizează un lanţ Markov continuu, cu număr infinit de stări (cum este şi cazul procesului în discuţie), se poate face cu ajutorul transformatei Z (***funcţia generatoare de probabilităţi***):

$$G(z) = Z\{p_n\} = \sum_{n=0}^{\infty} p_n \cdot z^n, \text{ cu } |z| \le 1 \qquad (10.4.9)$$

Aplicaţia 10.4.1

Să se determine, aplicând transformata Z, probabilităţile staţionare de stare ale unui proces de naştere şi moarte caracterizat prin $\lambda_i = \lambda$ şi $\mu_i = \mu$, pentru orice $i > 0$.

Indicaţie: Se trasează diagrama ratelor de tranziţie corespunzătoare, pe baza căreia se scriu următoarele ecuaţii de echilibru global:

- pentru $j = 0$: $\lambda \cdot p_0 = \mu \cdot p_1$,
- pentru $j > 0$: $(\lambda + \mu) \cdot p_j = \lambda \cdot p_{j-1} + \mu \cdot p_{j+1}$

Cu scopul de a obţine serii infinite, de genul transformatei Z, egalităţile corespunzătoare ultimei relaţii se amplifică cu z^j şi se însumează, rezultând în final relaţia: $(\lambda+\mu) \cdot \underbrace{\sum_{j=1}^{\infty} p_j \cdot z^j}_{G(z)-p_0} = \lambda \cdot \underbrace{\sum_{j=1}^{\infty} p_{j-1} \cdot z^j}_{z \cdot G(z)} + \mu \cdot \underbrace{\sum_{j=1}^{\infty} p_{j+1} \cdot z^j}_{z^{-1} \cdot [G(z)-p_0-p_1 \cdot z]}$

Expresiile de sub acolade sunt deduse prin aplicarea proprietăţilor următoare:

$$Z\{p_{n-j}\} = z^{+j} \cdot G(z)$$

$$Z\{p_{n+j}\} = z^{-j} \cdot \left(G(z) - p_0 - p_1 \cdot z - \ldots - p_{j-1} \cdot z^{j-1}\right)$$

care, în domeniul analizei semnalelor, unde exponentul variabilei z se consideră negativ, poartă denumirea de **întârziere**, respectiv de **avansare în timp** [Cons99], [MDS01]. Ţinând cont că $p_1 = \lambda/\mu \cdot p_0$, rezolvarea ecuaţiei în necunoscuta $G(z)$ conduce la următoarea soluţie:

$$G(z) = p_0/(1-\rho \cdot z), \text{ cu } \rho = \lambda/\mu.$$

Parametrul p_0 din relaţia precedentă se determină din relaţia de normare care, tradusă în planul z, se scrie: $G(z)\big|_{z=1} = 1$. Prin urmare, $p_0 = 1-\rho$ şi deci:

$$G(z) = (1-\rho)/(1-\rho \cdot z).$$

Determinarea *originalului*, deci a secvenţei de probabilităţi staţionare de stare p_j, se face, în general, aplicând transformata Z inversă. În cazul de faţă, este mult mai simplu să ţinem cont că, pentru $\rho < 1$, avem următorul rezultat:

$$(1-\rho \cdot z)^{-1} = \sum_{j=0}^{\infty} \rho^j \cdot z^j,$$

care, substituit în expresia lui $G(z)$, conduce la:

$$G(z) = \sum_{j=0}^{\infty} (1-\rho) \cdot \rho^j \cdot z^j,$$

cu ajutorul căreia, prin *identificarea termenilor,* ne oferă relaţia finală:

$$p_j = (1-\rho) \cdot \rho^j, \text{ cu } j \ge 0 \qquad (10.4.10)$$

capitolul 11

Traficul de telecomunicaţii

11.1 Definiţii, unităţi şi metode de măsurare

Traficul [Nic95] este o noţiune frecvent utilizată în limbajul cotidian, la care se face referire ori de câte ori dialogul are ca subiect o activitate mai mult sau mai puţin legală, ce presupune transferul printr-un anumit mijloc "a ceva anume" (obiecte, idei sau informaţii). Cu toată această frecventă uzanţă, ce ar da impresia unei înţelegeri clare a semnificaţiei, cuvântul trafic ascunde de fapt, din punct de vedere matematic, o serie de noţiuni nu tocmai simple, deduse şi introduse pentru a modela într-un mod riguros fenomenul corespunzător. Astfel, dacă considerăm procesul $N_a(t)$ al sosirilor şi X_i cantitatea asociată sosirii de identitate i, atunci:

$$S(t) = \sum_{i=1}^{N_a(t)} X_i \qquad (11.1.1)$$

reprezintă ***volumul de trafic*** (***traficul***) ***oferit*** până la momentul t, de către o anumită ***sursă de trafic*** (de exemplu, o firmă de produse alimentare) unui anumit ***sistem de servire*** (de exemplu, o firmă de transporturi auto), spre a fi transferat (de exemplu, către un anumit magazin alimentar) în condiţii ideale de servire, în care fiecare cerere de transfer este servită cu promptitudine (fără întârziere).

Conform definiţiei anterioare, mărimea $S(t)$ este un proces aleatoriu continuu, de sumare la momente repartizate aleatoriu, conform unui proces de numărare, $N_a(t)$ (figura 11.1.1). Termenii sumei reprezintă realizări particulare

generate de o serie X_i de variabile aleatorii independente şi identic distribuite, în mod generic notate X. În consecinţă, suma $S(t)$ prezintă o valoare medie, care este denumită ***trafic mediu***, fiind dată de relaţia (vezi aplicaţia 6.3.2.4):

$$\mathrm{E}[S(t)] = \mathrm{E}[N_a(t)] \cdot \mathrm{E}[X] \qquad (11.1.2)$$

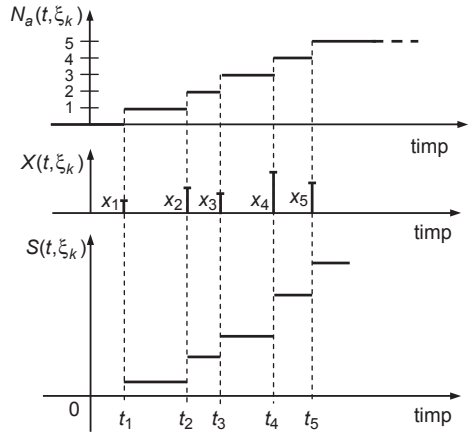


Figura 11.1.1: Realizări particulare de procese implicate în analiza traficului oferit (modelul 1)

În funcţie de domeniul considerat, cantităţile asociate sosirilor se pot măsura în kilograme, metri cubi, bucăţi etc. Dar, în majoritatea situaţiilor, inclusiv cel al reţelelor de telecomunicaţii, aceste cantităţi nu sunt recepţionate, respectiv transferate instantaneu. În consecinţă, fiecărei sosiri, numită ***cerere de serviciu***, îi sunt asociate, în general, un debit (m^3/oră, biţi/sec, pachete/sec etc.) sau o bandă (audio, video etc.) şi un ***timp de servire*** (cum este cazul apelurilor telefonice sau transferurilor de fişiere). În această situaţie (figura 11.1.2), ***volumul de trafic*** $S(t)$ ***oferit*** până la momentul nu poate lua în considerare şi duratele care depăşesc această valoare (adică t). Prin urmare, *traficul mediu* este dat de relaţia:

$$\mathrm{E}[S(t)] = \mathrm{E}[N_a(t)] \cdot \mathrm{E}[X'] \qquad (11.1.3)$$

în care:

$$X' = \begin{cases} X & (X \le t-a) \\ t-a & (X > t-a) \end{cases} \qquad (11.1.4)$$

unde a este momentul sosirii, cu valorile particulare a_1, a_2, a_3 şi a_4 în figura 11.1.2, ce poate fi plasat oriunde în intervalul $(0,t)$.

Figura 11.1.2 pune în evidenţă o altă mărime ce decurge din trafic, fiind derivată la propriu din acesta, şi anume ***intensitatea*** (instantanee a) ***traficului***, care este definită prin relaţia:

$$A(t) = \mathrm{d}S(t)/\mathrm{d}t \qquad (11.1.5)$$

şi care este, de fapt, *numărul de clienţi*, $N(t)$, aflaţi în sistem la momentul t.

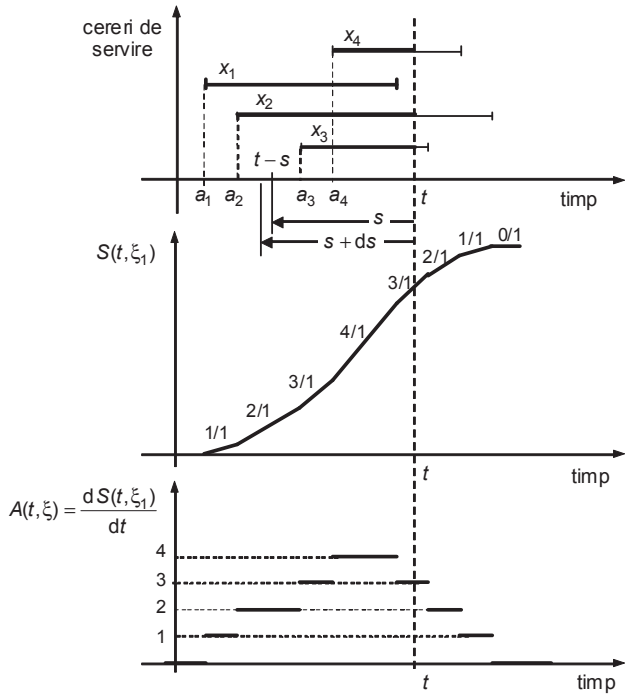


Figura 11.1.2: Realizări particulare de procese implicate în analiza traficului (modelul 2)

Desigur şi această mărime are o medie, care este numită ***intensitate medie*** (instantanee) ***a traficului*** şi este dată de relaţia:

$$\mathrm{E}\left[A(t)\right] = \frac{\mathrm{dE}\left[S(t)\right]}{\mathrm{d}t} = \frac{\mathrm{d}\left(\mathrm{E}\left[N_a(t)\right]\cdot\mathrm{E}\left[X'\right]\right)}{\mathrm{d}t} \qquad (11.1.6)$$

Relaţia (11.1.6) nu este unică prin care se poate calcula intensitatea medie instantanee a traficului. Astfel, dacă considerăm $N_a(t)$ procesul de sosire

al clienţilor ($N_a(t) = 0$, $t \le 0$) şi $N_d(t)$ procesul plecărilor din sistem (în urma satisfacerii serviciilor cerute), atunci intensitatea medie instantanee a traficului, fiind totuna cu numărul mediu instantaneu de clienţi în sistem poate fi exprimată prin următoarea diferenţă:

$$\mathrm{E}\left[A(t)\right] = \mathrm{E}\left[N_a(t)\right] - \mathrm{E}\left[N_d(t)\right] \qquad (11.1.7)$$

Numărul mediu de plecări în intervalul $(0,t)$ este dat de relaţia:

$$\mathrm{E}\left[N_d(t)\right] = \int_0^\infty P\left[X \le s\right]\cdot\left\{-\mathrm{d}\,\mathrm{E}\left[N_a(t-s)\right]\right\} \qquad (11.1.8)$$

în care, expresia de integrat reprezintă numărul mediu de clienţi care intră în sistem în intervalul $(t-s-\mathrm{d}s, t-s)$ **şi** care îl părăsesc înainte de momentul t, adică produsul dintre variaţia, în sens invers axei timpului (vezi figura 11.1.2) a numărului mediu de clienţi care intră în sistem în intervalul amintit (al doilea termen) **şi** "procentul" clienţilor care părăsesc sistemul înaintea expirării duratei *s* (primul termen), adică până la momentul t.

Integrala cuprinsă în relaţia (11.1.8) poate fi transformată aplicând metoda integrării prin părţi. În consecinţă, se obţine:

$$\mathrm{E}\left[N_d(t)\right] = \int_0^\infty \mathrm{E}\left[N_a(t-s)\right]\cdot\mathrm{d}P\left[X \le s\right] = \int_0^\infty \mathrm{E}\left[N_a(t-s)\right]\cdot f_X(s)\cdot\mathrm{d}s \qquad (11.1.9)$$

ceea ce permite rescrierea relaţiei (11.1.7) sub forma:

$$\mathrm{E}\left[A(t)\right] = \int_0^\infty \mathrm{E}\left[N_a(t) - N_a(t-s)\right]\cdot f_X(s)\,\mathrm{d}s \qquad (11.1.10)$$

Această ultimă relaţie permite o interpretare diferită a intensităţii medii instantanee a traficului decât cea a numărului de clienţi în sistem şi anume ca număr mediu de clienţi care sosesc pe durata medie a unui timp de servire (convorbire în cazul apelurilor telefonice).

Pe lângă relaţiile (11.1.6) şi (11.1.10), intensitatea medie instantanee a traficului se mai poate stabili şi plecând de **frecvenţa** (**rata**) **medie *instantanee* a sosirilor**, noţiune definită astfel:

$$\lambda(t) = \mathrm{d}\,\mathrm{E}\left[N_a(t)\right]/\mathrm{d}t \qquad (11.1.11)$$

În consecinţă, $\lambda(t-s)\cdot\mathrm{d}s$ reprezintă numărul mediu de apeluri sosite pe un interval infinitezimal, $\mathrm{d}s$, începând cu momentul $t-s$, iar $\lambda(t-s)\cdot P\left[X > s\right]\cdot\mathrm{d}s$ reprezintă numărul mediu de apeluri sosite pe intervalul $\mathrm{d}s$, începând cu momentul $t-s$ şi care mai sunt active la momentul t. Deci, aplicând teorema probabilităţii totale, se ajunge la a treia relaţie de calcul pentru intensitatea instantanee a traficului, adică:

$$\mathrm{E}\left[A(t)\right] = \int_0^t \lambda(t-s)\cdot\left(1 - F_X(s)\right)\cdot\mathrm{d}s \qquad (11.1.12)$$

În general, ceea ce interesează pentru o caracterizare globală a fenomenului este valoarea pe care o ia intensitatea medie instantanee a traficului

după un timp suficient de mare, adică ***intensitatea medie a traficului***, notată A (sau ρ) şi dată de expresia:

$$A = \lim_{t\to\infty} \mathrm{E}\big[A(t)\big] \qquad (11.1.13)$$

În aceste condiţii ($t \to \infty$), volumul mediu de trafic corespunzător timpilor de servire, contorizaţi dincolo de momentul t, al unor cereri iniţiate înainte de acest moment, devine neglijabil faţă de volumul de trafic acumulat până în acel moment. Prin urmare, "surplusul" fiind neglijabil, putem folosi pentru calculul volumului mediu de trafic acumulat după un timp suficient de mare, indiferent cum se însumează variabilele X_i, relaţia (11.1.2) modificată astfel:

$$\lim_{t\to\infty} \mathrm{E}\big[S(t)\big] = \lim_{t\to\infty} \mathrm{E}\big[N_a(t)\big]\cdot \mathrm{E}\big[X\big] \qquad (11.1.14)$$

Intensitatea medie a traficului se poate calcula şi pe baza relaţiei:

$$A = \lim_{t\to\infty} \frac{S(t)}{t} \qquad (11.1.15)$$

Fracţia conţinută în relaţia (11.1.15) poate fi pusă sub forma:

$$\frac{S(t)}{t} = \frac{N_a(t)}{t}\left\{\frac{1}{N_a(t)}\sum_{i=1}^{N_a(t)} X_i\right\} \qquad (11.1.16)$$

care, în final, ţinând cont de relaţia (6.6.13), conduce la următorul rezultat:

$$A = \frac{\mathrm{E}[X]}{\mathrm{E}[T]} = \lambda\cdot \mathrm{E}\big[X\big] \qquad (11.1.17)$$

unde: $T =$ timpul dintre două sosiri succesive, iar

$\lambda =$ frecvenţa (rata) medie a sosirilor.

Dat fiind utilizare sa frecventă în raport cu celelalte noţiuni din care derivă, intensitatea medie a traficului se denumeşte, simplu, ***trafic***. Desigur că această referire abuzivă poate crea ambiguităţi, dar ele sunt îndepărtate analizând "expresia" acestei mărimi, care, dacă este o constantă, reprezintă, cu siguranţă, o intensitate medie de trafic.

Aplicaţia 11.1.1

Determinaţi caracteristicile traficului relative la intensitatea sa, în ipoteza că sosirile urmează un proces de rată λ, iar timpii de servire o distribuţie exponenţială de medie $1/\mu$.

Rezolvare:

- **Varianta I:** $\mathrm{E}\big[X_a\big] = \int_0^{t-a} x\cdot f_X(x)\,\mathrm{d}x + \int_{t-a}^{\infty} (t-a)\cdot f_X(x)\,\mathrm{d}x = \frac{1}{\mu}\cdot\big[1-\mathrm{e}^{-\mu\cdot(t-a)}\big]$

reprezintă media până la momentul t, a duratei unui apel iniţiat la momentul a.

Deci, indiferent de moment: $\mathrm{E}\big[X'\big] = \frac{1}{t}\int_0^t \mathrm{E}\big[X_a\big]\,\mathrm{d}a = \frac{1}{\mu}\left(1+\frac{1}{t\cdot\mu}\mathrm{e}^{-\mu\cdot t}\right)$.

Prin urmare, conform relaţiilor (11.1.3) şi (11.1.6), rezultă că:

$$\mathrm{E}\big[S(t)\big] = \frac{\lambda}{\mu}t + \frac{\lambda}{\mu^2}\mathrm{e}^{-\mu\cdot t} \underset{t\to\infty}{\to} \frac{\lambda}{\mu}t \quad \text{şi} \quad \mathrm{E}\big[A(t)\big] = \frac{\lambda}{\mu}\left(1-\mathrm{e}^{-\mu\cdot t}\right) \underset{t\to\infty}{\to} \frac{\lambda}{\mu} = A$$

- **Varianta II:** se aplică relaţia (11.1.10), deci:

$$\mathrm{E}\big[A(t)\big] = \int_0^{\infty}\Big\{\mathrm{E}\big[N_a(t)\big]-\mathrm{E}\big[N_a(t-s)\big]\Big\}\cdot f_X(s)\cdot\mathrm{d}s$$

$$= \int_0^t\big[\lambda\cdot t-\lambda\cdot(t-s)\big]\cdot\mu\,\mathrm{e}^{-\mu\cdot s}\cdot\mathrm{d}s = \cdots = \lambda/\mu\cdot(1-\mathrm{e}^{-\mu\cdot t}) \underset{t\to\infty}{\to} \lambda/\mu = A$$

- **Varianta III:** considerând relaţia (11.1.12) se obţine:

$$\mathrm{E}\big[A(t)\big] = \int_0^t \lambda\,\mathrm{e}^{-\mu\cdot s}\mathrm{d}s = \lambda/\mu\cdot\left(1-\mathrm{e}^{-\mu\cdot t}\right) \underset{t\to\infty}{\to} \lambda/\mu = A$$

- **Varianta IV:** se utilizează relaţia (11.1.17) cu $\mathrm{E}[X] = 1/\mu$ şi $\mathrm{E}[T] = 1/\lambda$.

Rezultă astfel că: $A = \frac{1/\mu}{1/\lambda} = \frac{\lambda}{\mu}$

* * *

Dat fiind faptul că în cazul serviciilor de telefonie traficul se raportează la durate ale convorbirilor, unitatea sa de măsurare este adimensională. CCITT a decis în 1964 ca unitatea internaţională de trafic să poarte denumirea de ***ERLANG***, în cinstea matematicianului danez A.K. Erlang, ca o recunoaştere a cercetărilor sale fundamentale în domeniul traficului.

Conform definiţiei CCITT, azi ITU-T

1 Erlang reprezintă traficul de ocupare continuă a unui circuit sau dispozitiv sau grup înlănţuit de circuite sau dispozitive în decurs de o oră.

În practică, se mai utilizează însă şi alte unităţi de trafic, ce sunt derivate din diferite durate etalon de apel (2 minute, 100 secunde sau 1 minut), şi anume:

- **EBCH** (***E**quated **B**usy **H**our **C**all*) = **ARHC** (***A**ppel **R**éduit à l'**H**eure **C**hargée*);
- **CCS** (***C**ent **C**all **S**econds*);
- **CM** (***C**ommunication **M**inute*).

Între toate aceste unităţi există următoarele relaţii de echivalenţă:

$$1\,\mathrm{E} = 36\,\mathrm{CCS} = 30\,\mathrm{ARHC} = 30\,\mathrm{EDCH} = 60\,\mathrm{CM} \qquad (11.1.18)$$

În cazul reţelelor şi a serviciilor de date, volumul de trafic se exprimă în multipli de octeţi (Bytes), iar intensitatea traficului în multiplii de octeţi pe secundă. În consecinţă, traducerea în unităţi Erlang a unui trafic măsurat în astfel de unităţi se face luând în considerare capacitatea de servire a respectivului

trafic, ceea ce permite *conversia* cantităţii specificate în durate corespunzătoare, de servire.

Aplicaţia 11.1.2

Calculaţi traficul şi volumul mediu de trafic oferit timp de 10 ore de către o sursă de trafic, care generează blocuri informaţionale de lungime medie 1200 octeţi cu o rată medie de 100 bloc/sec, unui sistem înzestrat cu o capacitate de servire de 2,048 Mbit/sec. Rezultatele se vor exprima în ambele unităţi de măsură.

* * *

Sistemele şi reţelele de telecomunicaţii constituie, în principal, suportul de "desfăşurare" a traficului informaţional dintre utilizatorii acestuia. Din acest motiv, dimensionarea şi dezvoltarea lor se bazează pe cunoaşterea caracteristicilor serviciilor şi cererilor de serviciu cărora le sunt destinate.

Cunoaşterea caracteristicilor se realizează şi se reactualizează continuu printr-o ***analiză de trafic***, realizată periodică şi care presupune achiziţia de date şi prelucrarea acestora în diverse moduri. Astfel, considerând T durata de observare (înregistrare) şi R numărul de resurse considerate, în practică se pot aplică următoarele metode:

❖ **Metoda I**: notăm cu $\sigma(i)$ durata cumulată de servire în care i resurse sunt ocupate simultan. În aceste condiţii, înseamnă că estimatorul ***traficului scurs*** pe o durată T se exprimă prin relaţia: $\hat{S} = \sum_{i=0}^{R} i \cdot \sigma(i)$. Deci estimatorul intensităţii traficului scurs este:

$$\hat{A} = \frac{\hat{S}}{T} = \frac{1}{T}\sum_{i=0}^{R} i \cdot \sigma(i) \qquad (11.1.19)$$

❖ **Metoda II**: notăm cu σ_j durata cumulată de ocupare a resursei de identitate j, deci mărimile anterior definite se exprimă acum prin relaţiile:

$$\hat{S} = \sum_{j=1}^{R} \sigma_j \text{ şi } \hat{A} = \frac{1}{T}\sum_{j=1}^{R} \sigma_j \qquad (11.1.20)$$

❖ **Metoda III**: notăm cu n_j numărul de angajări şi cu σ_j durata cumulată de ocupare a resursei j pe durata T şi calculăm:

i) estimatorul timpului mediu de servire:

$$\hat{\tau}_j = \sigma_j / n_j \qquad (11.1.21)$$

ii) estimatorul ratei de angajare:

$$\hat{\lambda}_j = n_j / T \qquad (11.1.22)$$

pe baza cărora se obţine că:

$$\hat{A} = \sum_{j=1}^{R} \hat{\lambda}_j \cdot \hat{\tau}_j \qquad (11.1.23)$$

Aplicaţia 11.1.3

În urma observării unui grup de 4 resurse pe durata a 6 minute s-au trasat diagrame de utilizare corespunzătoare, ca în figura 11.1.3, în care simbolurile 1 şi 0 reprezintă starea de ocupare, respectiv de neutilizare a fiecărei resurse.

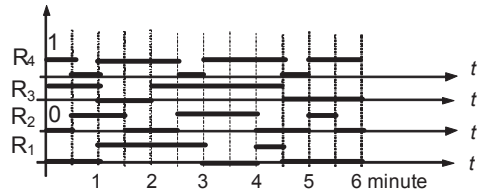


Figura 11.1.3: Diagramă exemplu de utilizare resurse

Stabiliţi intensitatea medie a traficului servit de respectivul grup pe parcursul intervalului considerat.

Rezolvare: mărimile de lucru ale celor trei metode prezentate iau valorile din tabelele 11.1.1 şi 11.1.2.

Calculele finale conduc, desigur, la aceleaşi valori, adică pentru o perioadă de observaţie de 6 minute : $\hat{S} = 13{,}5$ minute, deci $\hat{A} = 2{,}25$ E.

Tabelul 11.1.1

i	$\sigma(i)$ [min]	$i \cdot \sigma(i)$
0	0,5	0,0
1	0,5	0,5
2	2,0	4,0
3	3,0	9,0
4	0	0,0
Total		13,5 min

Tabelul 11.1.2

j	n_j	σ_j [min]	$\hat{\lambda}_j$ [min^{-1}]	$\hat{\tau}_j$ [min]	$\hat{A}_j$
1	2	2,5	1/3	5/4	5/12
2	3	3,0	1/2	1	1/2
3	2	3,5	1/3	7/4	7/12
4	4	4,5	2/3	9/8	3/4
Total	11	13,5	11/6	27/22	9/4

* * *

Cele trei metode prezentate au, în principal, ca obiectiv determinarea unei valori medii(cea a traficului), dar rezultatele intermediare oferite de acestea se pot utiliza şi pentru elaborarea unei analize mai profunde, cu scopul stabilirii unor distribuţii empirice de interes. În acest sens, se pot determina, de exemplu, ***frecvenţele relative*** pentru diverse evenimente cum ar fi, de exemplu:

$$E = \{i \text{ resurse sunt simultan ocupate}\}.$$

şi pentru care formula de calcul a frecvenţei relative este:

$$f_T(E) = \sigma(i)/T \qquad (11.1.24)$$

Luând în considerare exemplul considerat în aplicaţia anterioară, 11.1.3, formula (11.1.24) conduce la rezultatele conţinute în tabelul 11.1.3; pe baza acestora s-au trasat reprezentările din figura 11.1.4, care precizează probabilităţile elementare empirice (a), şi respectiv funcţia de distribuţie empirică corespunzătoare (b).

Tabelul 11.1.3

i	**0**	**1**	**2**	**3**	**4**
$\sigma(i)/T$	1/12	1/12	4/12	6/12	0

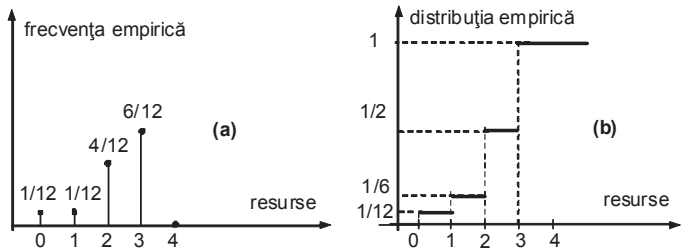


Figura 11.1.4: Exemple de frecvenţe şi distribuţii empirice

Din analiza acestor două reprezentări se poate deduce, de exemplu, că cele mai frecvente situaţii în care se poate găsi sistemul sunt cele cu 2 sau 3 resurse ocupate. De asemenea, rezultatele obţinute evidenţiază faptul că sistemul considerat este supradimensionat, ceea ce poate constitui un avantaj atât în prezent prin redundanţa asigurată, cât şi în viitor, când este de aşteptat o creştere a frecvenţei cererilor de servicii.

11.2 Surse de trafic

Traficul deservit de reţelele de telecomunicaţii poate avea atât o reprezentare analogică (este cazul semnalelor audio sau video), cât şi numerică (cazul transmisiilor de date, al semnalizărilor între centrale etc.). Clasificarea semnalelor nu este însă foarte restrictivă şi definitivă, deoarece prin codare semnalele audio şi video pot deveni numerice (vezi cazul transmisiunilor multiplexate în timp), iar cele numerice devin analogice prin modulare (vezi cazul modem-urilor). În consecinţă, mijloacele de transfer (sisteme de transmisiuni, noduri de comutaţie) trebuie să utilizeze tehnici şi tehnologii de natură corespunzătoare.

Dat fiind avantajele tehnicilor şi tehnologiilor numerice, tendinţele actuale manifestate în cadrul reţelelor de telecomunicaţii sunt numerizarea completă şi integrarea tuturor serviciilor. În principiu, ***integrarea*** permite implementarea unei

reţele unice care este capabilă să transfere orice tip de trafic. În aceste condiţii, varianta de clasificare a surselor de trafic, prezentată în continuare, este cea utilizată în cadrul descrierii reţelelor de bandă largă, care au adoptat modul de transfer asincron, [Mrt91], **ATM** (***A**synchronous **T**ransfer **M**ode*), pentru satisfacerea unei varietăţi ample de servicii.

11.2.1. Clasificare

Principalele criterii care stau la baza clasificării surselor de trafic, indiferent de domeniu, sunt caracteristicile şi cerinţele de calitate a servirii, ***QoS*** (***Q**uality **o**f **S**ervice*), manifestate de către acestea. Astfel, din specificaţiile oferite de "Forumul ATM", organism destinat reglementărilor în domeniul reţelelor ATM, se desprind următoarele patru ***clase*** (*de surse*) ***de trafic*** (***servicii***):

- **Clasa A** – surse cu debit constant, **CBR** (***C**onstant **B**it **R**ate*), şi restricţii temporale severe, de genul voce sau imagine fără compresie;
- **Clasa B** – surse cu debit variabil şi restricţii temporale severe, **RT-VBR** (***R**eal **T**ime **V**ariable **B**it **R**ate*), de genul voce sau imagine cu compresie;
- **Clasa C** – surse cu debit variabil şi restricţii temporale indulgente, **NRT-VBR** (***N**on **R**eal **T**ime **V**ariable **B**it **R**ate*), de genul tranzacţiilor în timp util, precum rezervarea unui loc într-o cursă aeriană;
- **Clasa D** – surse cu debit variabil şi fără restricţii temporale, **VBR** (***V**ariable **B**it **R**ate*), de genul poştă electronică sau transfer de fişiere.

Clasificarea sus menţionată nu este unică, pe lângă acesta fiind demnă de menţionată varianta următoare, care are în vedere şi contextul reţelelor celulare. Şi în acest caz, clasele de servicii sunt în număr de 4, denumirile şi caracteristicile lor, specifice, fiind următoarele:

1. clasa serviciilor conversaţionale – se referă la comunicaţiile bidirecţionale, în timp real, de genul convorbirilor telefonice, fiind, prin urmare, extrem de **sensibile la întârzieri** şi mai puţin la pierderi a căror limită este impusă de calitatea acceptată de utilizatori, privind fidelitatea cu care semnalele sunt reproduse la destinaţie. În plus, serviciul este simetric, în ceea ce priveşte traficul pe **legătura ascendentă UL** (***U**p **L**ink*) şi pe **legătura descendentă DL** (***D**own **L**ink*), dat fiind poziţia, de egalitate, a interlocutorilor.

2. clasa serviciilor tip *streaminig* – se referă la fluxurile de date, de genul **difuziune audio** sau **video**, a căror procesare, la destinaţie, în vederea reproducerii informaţiei, se desfăşoară în timp real. Astfel, se evită întârzierile semnificative, ce apar dacă întregul conţinut trebuie, mai întâi, să fie descărcat (downloaded) şi se permite sursei să schimbe dinamic (în timpul emisiei) conţinutul fluxului (precum în cazul programelor radio în direct). Fată de cazul precedent, cerinţele privind întârzierile sunt mai puţin severe, pierderile sunt acceptate în aceleaşi limite, iar traficul pe cele două direcţii, este asimetric, ceea mai semnificativă componentă întâlnindu-se pe legătura descendentă.

3. clasa serviciilor interactive – se referă la transferurile sigure de date corespunzătoare aplicaţiilor de genul motoare de căutare (*web browsing*) sau accesări de servere (*server access*). În acest caz, vorbim de o **sensibilitate** extremă **la pierderi**, întârzierile maxime acceptate depinzând de temporizările impuse de protocoalele utilizate, de comunicaţie (HTTP, FTP, TCP etc.). La acestea se adaugă *productivitatea* (*throughput*) a cărei valoare trebuie să satisfacă aşteptările utilizatorilor, privind acest gen de servicii.

4. clasa serviciilor de fundal (***background***) – se referă la transferurile de date care tolerează întârzieri mari, debite mici şi au cea mai scăzută prioritate, unica cerinţă fiind integritatea informaţiei, precum în cazul aplicaţiei *email*.

În principiu, o sursă de trafic prezintă, pentru fiecare din nivelele 1 – 5 ale modelului OSI, anumite caracteristici cum ar fi, în cazul nivelului sesiune, distribuţia duratei de menţinere a unei conexiuni. Dintre aceste caracteristici, cele mai des referite, luate în considerare în cazul primei clasificării prezentate anterior, se raportează la nivelul reţea şi, în conformitate cu figura 11.2.1.1, sunt:

- ***Debitul* (*rata*) *de vârf,*** **PIR** (***P**eak **I**nformation **R**ate*) – este rata maximă de generare a traficului pentru o conexiune (evident în timpul perioadelor sale de activitate). Conform recomandărilor ITU, aferente tehnologiei ATM, au fost introduse două definiţii, şi anume:
  - ***Debitul de vârf instantaneu*** se măsoară în celule/sec. şi este egal cu inversul timpului minim între două generări succesive de celule, **3-PDU** (***P**rotocol **D**ata **U**nit* de nivel ***3***, pentru reţele ATM);
  - ***Debitul de vârf integrat*** este numărul mediu de celule generate într-un interval, T_a, de activitate intensă, raportat la acesta.
- ***Debitul* (*rata*) *medie*** – definit şi acesta în două moduri, şi anume:
  - ***Debitul mediu real*** este numărul de celule generate de o sursă pe întreaga durată a conexiunii, raportat la aceasta;
  - ***Debitul mediu*** este numărul de celule generate de o sursă pe o durată, W, suficient de mare, raportată la aceasta.

În cazul reţelelor IP, unde blocurile informaţionale tranzitate, în speţă: pachete, au lungimi variabile, mărimile de mai sus au următoarea semnificaţie:

- ***Debitul de vârf instantaneu*** se măsoară în octeţi/sec. şi reprezintă viteza maximă cu care poate fi transmisă informaţia prin stratul reţea;
- ***Debitul de vârf integrat*** este numărul mediu de octeţi generaţi într-un interval, T_a, de activitate intensă, raportat la acesta;
- ***Debitul mediu real*** este numărul de octeţi generaţi de o sursă pe întreaga durată a conexiunii, raportat la aceasta;
- ***Debitul mediu*** este numărul de octeţi generaţi de o sursă pe o durată, W, suficient de mare, raportată la aceasta.

- ***Perioada de activitate*** – intervalul de timp în care traficul (pachete, celule) generat de o sursă este semnificativ.
- ***Perioada de inactivitate* (*tăcere*)** - se poate defini ca intervalul de timp între finalul unei perioade de activitate şi începutul următoarei perioade de activitate, în care traficul generat de o sursă este nesemnificativ.

O sursă de trafic se consideră activă dacă generarea de celule (pachete) se face în condiţii normale sau intense. Momentul când o sursă devine activă sau inactivă se poate stabili prin compararea duratei dintre două generări succesive de celule (pachete) cu o valoare de prag, t_p. Astfel, dacă pragul este depăşit (figura 11.2.1.1), se consideră că sursa trece în inactivitate sau desfăşoară o activitate sporadică. Revenirea la activitate coincide desigur cu momentul apariţiei primei celule (pachet) care, faţă de precedenta, se află la un interval sub pragul t_p.

- ***Coeficient de rafală*** (*sporadicity* sau *burstiness*) – este raportul dintre rata de vârf şi rata medie.

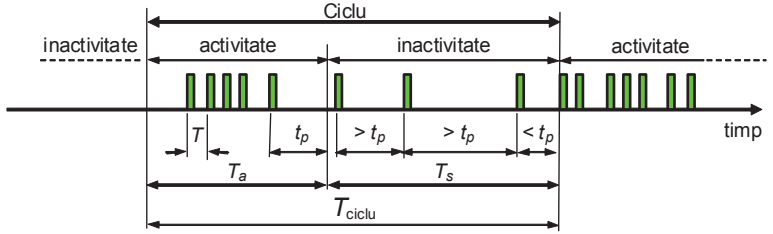


Figura 11.2.1.1: Înlănţuirea stărilor în cazul surselor ATM cu debit variabil

În funcţie de valoarea instantanee a debitului, faţă de mărimile precizate anterior, o sursă de trafic se poate găsi, în general:

- pe parcursul *perioadei de inactivitate*, în stările:
  - inactivă,
  - cu generare sporadică,
- pe parcursul *perioadei de activitate*, în stările:
  - cu generare normală (medie),
  - cu generare intensă (de vârf).

În concluzie, caracterizarea surselor cu debit variabil se face în principiu prin intermediul următorilor *parametri*:

- r_p = debitul de vârf instantaneu,
- r_m = debitul mediu,
- $\mathrm{E}[B]$ = lungimea medie a pachetelor,

- $\mathrm{E}[N_a]$ = numărul mediu de celule (pachete) generate într-o perioade de activitate,
- $\mathrm{E}[T_{\text{ciclu}}]$ = timpul mediu între începutul unei perioade de activitate şi începutul următoarei perioade de activitate,
- β = coeficientul de rafală,
- a^{-1} = durata medie a perioadei de activitate,
- s^{-1} = durata medie a perioadei de inactivitate (***s**ilence*),
- α = **frecvenţa activităţii** în cadrul conexiunii, numită şi **factorul de activitate** (**activity factor**)
- $\mathrm{E}[T]$ = intervalul mediu între două generări succesive de celule (pachete), în perioadele de activitate.

Între toţi aceşti parametri există relaţiile ce urmează:

$$\begin{aligned}
&\mathrm{E}[T_{\text{ciclu}}] = a^{-1} + s^{-1}; \quad \alpha = a^{-1}/(a^{-1} + s^{-1});\\
&\mathrm{E}[N_a] = a^{-1}/\mathrm{E}[T];\\
&r_m = \alpha/\mathrm{E}[T];\\
&r_p = \min\{T\} \text{ pentru reţele ATM};\\
&r_m = \alpha \cdot \mathrm{E}[B]/\mathrm{E}[T] \text{ pentru IP};\\
&\beta = r_p/r_m .
\end{aligned} \qquad (11.2.1)$$

11.2.2. Modelarea surselor de trafic

Necesitatea dezvoltării de mijloace analitice şi de programe de simulare, specializate pentru activităţile de proiectare, analiză şi optimizare a reţelelor de telecomunicaţii, a condus la identificarea mai multor modele matematice, care, funcţie de ipotezele asumate, aproximează în bună măsură sursele reale de trafic. În principiu, modelele probabilistice elementare, posibil de utilizat în descrierea aproximativă a activităţii unei surse de trafic sunt distribuţiile Bernoulli, geometrică, binomială, Poisson şi Gauss. Aceste modele nu sunt însă satisfăcătoare atunci când descrierea se referă la surse sofisticate de trafic. Din acest motiv, studiile teoretice şi experimentale au generat noi modele, [Mrt91], [Frost94], cu grad sporit de complexitate, dintre care cele mai reprezentative urmează a fi prezentate în continuare.

Contextul real, la care se referă prezentarea ce urmează, este înfăţişat în figura 11.2.2.1. Cele patru diagrame temporale iau în considerare nivelele la care se face caracterizarea fenomenului aleatoriu observat, propriu transferurilor de informaţie prin reţelele cu comutaţie de pachete, acestea fiind:

- **nivelul sesiune,**
- **nivelul rafală** (o rafală corespunde unei perioade de activitate din cadrul unei sesiuni, în care are loc transfer de informaţie),
- **nivelul pachet,**
- **nivelul celulă.**

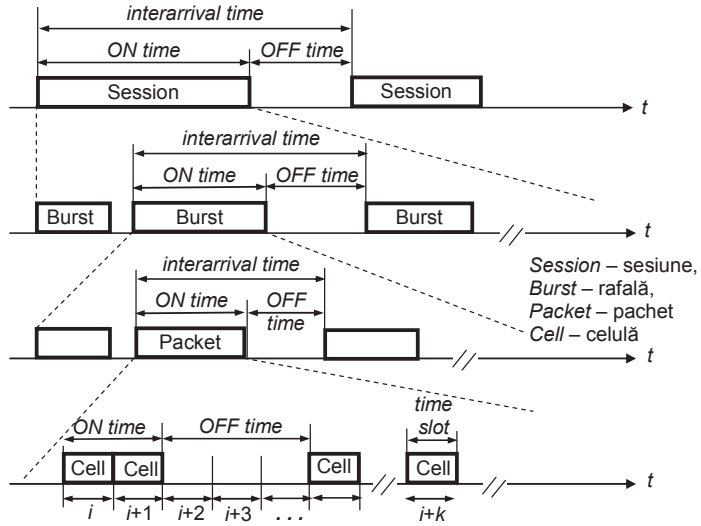


Figura 11.2.2.1: Nivele de detaliu considerate în procesul de modelare a surselor de trafic

Figura 11.2.2.1 mai conţine câteva mărimi precizate în engleză, dat fiind larga lor referire în oricare altă limbă, a căror semnificaţie este:

- ***interarrival time*** – timpul dintre două sosiri succesive (de sesiuni, rafale, pachete sau celule),
- ***ON time*** – timpul de activitate (pe diversele nivele de analiză),
- ***OFF time*** – timpul de inactivitate (pe diversele nivele de analiză),
- ***time slot*** – intervalul temporal care duce la discretizarea axei timpului în cazul în care analiza are în vedere nivelul celulă.

11.2.2.1. Surse de trafic ON/OFF

Multe din sursele de trafic actuale nu mai pot fi caracterizate la nivel de debit prin intermediul unei anumite distribuţii (binomială, gaussiană etc.). ***Sursele de trafic ON/OFF*** reprezintă un asemenea de caz: traficul emis *în rafale* de ele

constă dintr-o succesiune de *perioade de activitate,* ***ON***, în care generarea se face la debit constant (adică cu $\beta = 1$), şi *perioade de tăcere*, ***OFF***.

Figura 11.2.2.2 ilustrează o posibilă realizare a unui proces ce reprezintă debitul generat de o sursă ON/OFF, în care:

- A reprezintă debitul de vârf;
- u.i. este o unitate de informaţie (bit, celulă, cadru etc.);
- u.t. este o unitate de timp (secundă, durată a unei celule etc.).

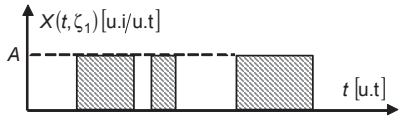


Figura 11.2.2.2: Exemplu de trafic generat de o sursă ON/OFF

Studiile experimentale au arătat că, în cazul surselor de trafic telefonic şi de date (cu anumite rezerve!), se poate considera că procesul $X(t)$ este staţionar şi că perioadele ON şi OFF urmează distribuţii exponenţiale. Aceste ipoteze permit modelarea matematică a surselor ON/OFF printr-un lanţ Markov a cărui diagramă de tranziţii se prezintă ca în figura 11.2.2.3, procesul fiind complet specificat de probabilităţile celor 2 stări:

$$p_0 = P\left[\{\text{Debit} = 0\}\right] = a/(a+s)$$
$$p_A = P\left[\{\text{Debit} = A\}\right] = s/(a+s) \overset{\text{notat}}{=} p \qquad (11.2.2.1)$$

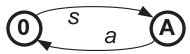


Figura 11.2.2.3: Diagrama de tranziţii pentru surse ON/OFF

Conform acestor relaţii, se poate aprecia că starea procesului reprezintă o variabilă aleatorie ce urmează o distribuţie Bernoulli de parametrul p. Prin urmare, debitul mediu al sursei este dat de relaţia:

$$D_{\text{mediu}} = p \cdot A \qquad (11.2.2.2)$$

În cazul în care se fac determinări experimentale, probabilitatea stării active se determină cu ajutorul formulei:

$$P\left[\{\text{Debit} = A\}\right] = \frac{D_{\text{mediu}}}{A} \qquad (11.2.2.3)$$

Tabelul 11.2.2.1 conţine valorile numerice ale parametrilor de trafic ce sunt caracteristice câtorva surse de trafic ON/OFF.

Tabelul 11.2.2.1: Parametri de trafic pentru surse ON/OFF

Tipul traficului	s [sec^{-1}]	a [sec^{-1}]	A [Mbit/sec]
Voce (semnal telefonic)	1 / 0,352	1 / 0,650	0,064
Imagine (fără schimbări bruşte)	3,9/(1 + 5,045[20])	3,9 – s	0,75 + 3,9 [20]
Date (transfer de fişiere)	1 / 0,1	1 / 1	2

În reţelele ATM serviciile CBR tip voce (telefonie) sunt transmise prin modulaţie MIC, deci rezultă un necesar de bandă de 64 kbit/sec.

În general, banda necesară pentru transmiterea unui semnal CBR audio este dată de formula:

$$\text{Banda audio} = (\text{număr biţi/eşantion}) \times (\text{număr eşantioane/sec}) \quad (11.2.2.5)$$

Astfel, pentru un semnal audio (de exemplu muzical) cu banda $10\,\text{Hz} \div 20\,\text{kHz}$, prin cuantizarea MIC cu 16 biţi/eşantion la o frecvenţă de eşantionare de $44{,}1\,\text{kHz}$, rezultă un necesar de bandă de $700\,\text{kbit/sec}$ mono, respectiv $1{,}4\,\text{Mbit/sec}$ stereo. Prin utilizarea unor tehnici adecvate de compresie necesarul de bandă pentru aplicaţia în discuţie poate fi redus până la $128 \div 196\,\text{kbit/sec}$. De asemenea, există posibilitatea transformării surselor video (VBR) în surse CBR, prin utilizarea unei scheme generale de tipul celei prezentate în figura 11.2.2.4.

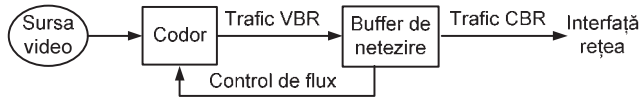


Figura 11.2.2.4: Sistem video CBR

Astfel, în urma codării informaţiei generate de o sursă video se obţine un trafic VBR, care este "nivelat" prin trecerea sa printr-un buffer de netezire. Pachetele (celulele ATM), din buffer-ul de netezire, sunt livrate reţelei în mod continuu (la intervale fixe de timp), ceea ce înseamnă că la intrarea în reţea traficul va fi de tip CBR. Dezavantajul major al acestei proceduri este *umplerea buffer-ului de netezire*, care este dimensionat în funcţie de întârzierea maxim admisibilă pentru traficul aferent unei aplicaţii.

Se poate însă remedia acest neajuns, prin intermediul unui mecanism de control al fluxului, implementat pe calea inversă: când gradul de umplere a buffer-ului de netezire depăşeşte un anumit prag, este "informat" codorul care va micşora rata proprie de transmisie. Consecinţa va fi însă şi o scădere a numărului de nivele de cuantizare, deci o degradare a calităţii serviciului.

În general, mecanismele de alocare de resurse în reţea pentru serviciile (sursele) continue sunt simple (pentru fiecare CBR se alocă o bandă egală cu rata de vârf a sursei CBR), dar coeficientul de utilizare a resurselor alocate pentru o sursa CBR poate avea valori foarte mici.

În tabelul 11.2.2.2 sunt prezentate câteva exemple de astfel de servicii, împreuna cu cerinţele aferente de bandă.

Tabel 11.2.2.2: Exemple de servicii cu trafic continuu (CBR)

Serviciul	Necesarul bandă [kbit/sec]
Telefonie	64
Stereo Hi-Fi	1400
Fax (grup 3)	14,4
Fax (grup 4)	64

11.2.2.2. Surse de trafic (continuu) variabil

Acest gen de surse generează trafic în mod continuu, dar debitul variază într-o anumită plajă de valori, după o distribuţie oarecare (ce poate fi chiar şi "normală"). Spre exemplificare se prezintă în continuare câteva categorii clasice de servicii cu trafic variabil.

- ***Serviciul VBR-video***

Constă într-o succesiune periodică de "cadre" care, datorită diferenţelor între cadrele consecutive precum şi a frecvenţei lor de succedare, permit ochiului uman să integreze aceste diferenţe într-o imagine în mişcare. Frecvenţa uzuală de succedare a cadrelor este de 24 sau 30 cadre/sec, numărul de biţi/pixel variind de la 8 (imagini monocrome) la 24 (imagini color – câte 8 biţi pentru fiecare din culorile elementare: roşu, verde, albastru), iar rezoluţia (număr pixeli/cadru) are valori cuprinse (tipic) între 128×120 pixeli/cadru (rezoluţie scăzută) la 640×480 pixeli/cadru (servere de studio – vizionare de filme) şi 1125 linii/cadru (HDTV). Banda necesară acestui tip de servicii se determină cu relaţia:

$$\text{Banda video} = \frac{\text{număr cadre}}{\text{secunde}} \times \frac{\text{număr pixeli}}{\text{cadru}} \times \frac{\text{număr biţi}}{\text{pixel}} \qquad (11.2.2.6)$$

Tabelul 11.2.2.3: Cerinţe de bandă pentru servicii VBR-video

SERVICII VIDEO	CERINŢE DE BANDĂ [kbit/sec]	
	Necompresat	Compresat
De timp real – 1/4 ecran – rezoluţie redusă (128×120 pixeli, 9 biti/pixel, 15 cadre/sec)	2.074	64
De timp real – 1/4 ecran – rezoluţie mare (128×240 pixeli, 9 biţi/pixel, 30 cadre/sec)	4.147	384
De timp real – ecran întreg (128×480 pixeli, 24 biţi/pixel, 30 cadre/sec)	8.294	2000
Server studio (filme) (640×480 pixeli, 24 biţi/pixel, 30 cadre/sec)	221.184	4000
HDTV(1125 linii, 24 biţi/pixel, 30 cadre/sec)	800.000	60.000 - 127.000

În tabelul 11.2.2.3 sunt prezentate câteva tipuri de servicii video cu cerinţele aferente de bandă.

- ***Servicii VBR-voce***

Spre deosebire de serviciile continue CBR-voce, serviciile discontinue VBR-voce nu transmit trafic decât în timpul perioadelor de activitate (nu se emite trafic în timpul cât sursa de voce este inactivă). Aceasta conduce la posibilitatea creşterii coeficientului de utilizare a resurselor alocate, resursele reţelei nemaifiind utilizate decât pentru transmisia informaţiei utile aferente unei surse de voce (perioada de activitate). În reţelele ATM, pentru conexiuni VBR-voce se utilizează o rată de vârf de 64 kbit/sec, astfel că o celulă ATM poate transporta $5{,}5 \div 5{,}875$ msec voce.

Prin observaţii experimentale, s-a constatat că perioadele de tăcere pentru surse cu debit variabil VBR-voce reprezintă circa $60 \div 65\,\%$ din timpul total de activitate al sursei (timpul cât durează comunicaţia propriu-zisă). De asemenea, s-a determinat o lungime medie a perioadei de activitate de $352\,\text{msec}$, respectiv $650\,\text{msec}$ pentru lungimea medie a perioadei de tăcere. Mai mult, modelele unanim acceptate astăzi pentru sursele cu trafic variabil de tip voce presupun ca duratele perioadelor de activitate şi respectiv tăcere au o distribuţie exponenţială.

- ***Servicii VBR-date***

Aceste servicii includ un spectru larg de comunicaţii, din punct de vedere al traficului putând avea caracteristici diferite (de exemplu un e-mail presupune transferul a câţiva kbiţi, iar într-un transfer de fişiere se pot transfera Mbiţi). Există categorii de aplicaţii ce necesită pentru transfer modul cu conexiune, altele se pot realiza în mod fără conexiune. Mai mult, este dificil de prezis a-priori cantitatea şi caracteristicile de trafic aferente unei conexiuni, chiar atunci când tipul aplicaţiei este cunoscut.

În cadrul aplicaţiilor de acest tip, ***interconectarea LAN-urilor prin ATM*** constituie un exemplu important, deoarece aici apar următoarele probleme:

- interconectarea LAN-urilor implică totdeauna comunicaţii între grupuri de utilizatori (ceea ce presupune realizarea unui serviciu fără conexiune), în timp ce ATM este orientat pe transferuri în mod circuit virtual;

- corespunzător tipului de LAN, pachetele pot avea lungimi diferite, iar ATM funcţionează cu celule de lungimi fixe (53 octeţi); deci pachetele trebuie segmentate în celule ATM la ieşirea din LAN-ul sursă şi reasamblate la intrarea LAN-ului destinaţie. În figura 11.2.2.5 este prezentată, spre exemplificare, distribuţia lungimii pachetelor pe reţelele Ethernet şi Token-Ring.

Lungimea pachetelor folosite depinde de tipul aplicaţiei, pentru reţelele Ethernet precizându-se că: pachetele lungi sunt utilizate în aplicaţii de tip transfer de fişiere, pachetele de lungime medie sunt utilizate în aplicaţii NFS, iar cele de lungime mică în comunicaţiile între terminal şi host.

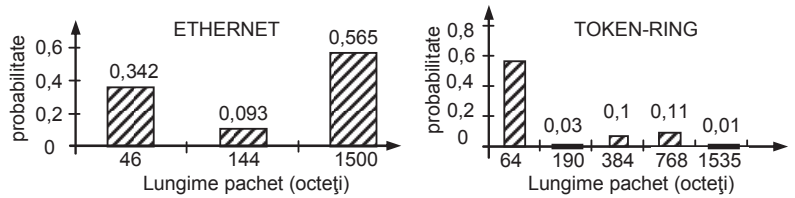


Figura 11.2.2.5: Distribuţia lungimii pachetelor în reţele LAN

În general, traficul caracteristic reţelelor Ethernet se poate modela ca un proces ON/OFF, cu lungime medie a pachetelor de 876,6 octeţi şi debit de vârf de 10 Mbit/sec. În ceea ce privesc reţelele tip Token-Ring, traficul se poate modela printr-un proces ON/OFF cu o medie a perioadei de activitate de 205 octeţi, cu coeficient de variaţie de 2,4 şi debit de vârf de 4 sau 16 Mbit/sec.

Un alt exemplu de servicii VBR-date îl constituie ***aplicaţiile multimedia*** utilizate în reprezentarea, stocarea, regăsirea şi transmiterea de informaţii prin medii multiple (conţinând text, voce, video, grafică, imagine, audio), de exemplu teleconferinţa, imagini medicale, reclame. În tabelul 11.2.2.4 se prezintă cerinţele de bandă pentru câteva aplicaţii tip multimedia.

Tabelul 11.2.2.4: Cerinţe de bandă pentru aplicaţii multimedia

APLICAŢIE	CERINŢA DE BANDĂ
- voce şi imagine de înaltă rezoluţie (diagnoză medicală etc.) - voce (32kbit/sec) şi o imagine/90sec (2000×2000/10)×12×1/90	85,3 kbit/sec
- teleconferinţă voce (32 kbit/sec) - 1,24 Moctet pagină text/60 sec (1,24×8/60 kbit/sec) - 35 koctet-grafică/60 sec (35×8/60 kbit/sec) - 2 imagini înaltă rezoluţie (60 koctet/60sec)	325 kbit/sec

De regulă, variaţia debitului generat de o sursă VBR se prezintă ca în figura 11.2.2.6.

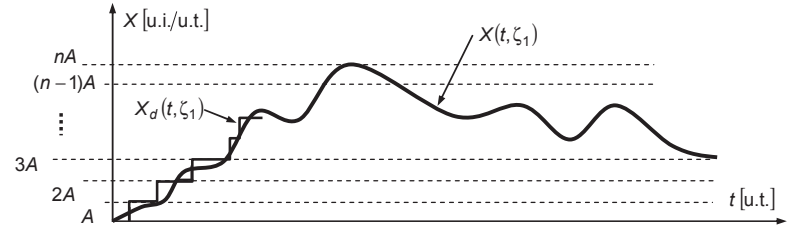


Figura 11.2.2.6: Cuantizarea unui debit continuu variabil

A. În cazul în care prezintă ***variaţii lente***, debitul unor astfel de surse poate fi aproximat de un debit discret obţinut prin cuantizarea liniară şi care este rezultatul superpoziţiei a *n* surse de trafic ON/OFF independente şi identic distribuite. Diagrama de tranziţii a procesului corespunzător este prezentată în figura 11.2.2.7. Probabilitatea de stare este în acest caz:

$$P\left[\{\text{Debit} = k \cdot A\}\right] = C_n^k \cdot p^k \cdot (1-p)^{n-k} \qquad (11.2.2.7)$$

Aproximarea prin discretizare este cu atât mai bună cu cât numărul paşilor este mai mare. Pentru $n > 20$, debitul generat de sursele de trafic cu variaţii lente poate fi considerat o variabilă aleatoare Gauss.

Figura 11.2.2.7: Diagrama de tranziţii pentru n surse ON/OFF cu trafic variabil lent

B. Debitele cu ***variaţii rapide*** pot fi şi ele aproximate prin debite discrete. În plus, dacă se poate considera că debitul discret rămâne pe fiecare nivel de cuantizare, pe o durată ce urmează o distribuţie exponenţială, atunci se alege ca model matematic un proces Markov, care prezintă şi tranziţii între stări neadiacente (figura 11.2.2.8).

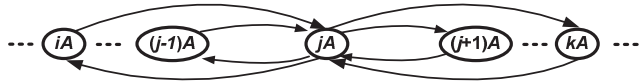


Figura 11.2.2.8: Diagrama de tranziţii pentru n surse ON/OFF cu trafic variabil rapid

Procesele prezentate mai sus, în care debitul instantaneu este modulat de un lanţ Markov continuu în timp, aparţin categoriei proceselor de tip **MMRP** (***M**arkov **M**odulated **R**ate **P**rocess*), cazul surselor ON-OFF fiind corespunzător proceselor **IRP** (***I**nterrupted **R**ate **P**rocess*). Dat fiind faptul că aceste procese nu aproximează satisfăcător toate genurile de surse reale de trafic, pe diverse nivele de studiu al acestora (serviciu, rafală, pachet/celulă, bit), studiile teoretice şi experimentale [Dob99] au avansat şi alte procese precum cele ce urmează:

❖ **Proces MMPP** (***M**arkov **M**odulated **P**oisson **P**rocess*) – un proces continuu de numărare care se desfăşoară conform unui proces Poisson a cărui rată medie, λ, se modifică sub controlul unui lanţ Markov continuu în timp.

❖ **Proces IPP** (***I**nterrupted **P**oisson **P**rocess*) – un caz particular al procesului MMPP, în care lanţul Markov modulator are numai 2 stări, incrementarea avansând doar într-una dintre ele.

❖ **Proces MMBP** (***M**arkov **M**odulated **B**ernoulli **P**rocess*) – un proces discret care se desfăşoară conform unui proces Bernoulli, al cărui parametru p (probabilitatea de apariţie) se modifică sub controlul unui lanţ Markov discret în timp.

❖ **Proces IBP** (***I**nterrupted **B**ernoulli **P**rocess*) – un caz particular al procesului MMBP, lanţul Markov modulator având numai 2 stări, pentru una din ele probabilitatea p fiind nulă.

În general, luând în considerare caracteristicile prezentate, procesele MMPP şi IPP sunt aplicate în cazul modelării surselor la nivel de pachet, iar procesele MMBP şi IBP sunt folosite când se modelează sursele la nivel de celulă (vezi figura 11.2.2.1). În privinţa acestor modele matematice, şi nu numai, următorul subcapitol vine cu informaţii suplimentare.

11.3.3.1. Model de amestec

Un astfel de model se obţine prin combinarea liniară a mai multor modele probabilistice, de complexitate inferioară, având, în consecinţă, o distribuţie, F, dată de relaţia:

$$F = \sum_{k=1}^{K} \lambda_k \cdot F_k \text{ , în care, pentru } k = \overline{1,K}: \qquad (11.3.7)$$

- F_k sunt distribuţiile modelelor incluse, iar
- λ_k sunt **coeficienţii de amestec** (adică "şansa" modelului M_k), legaţi între ei prin relaţia $\sum_k \lambda_k = 1$.

Aplicaţia 11.3.5

Reprezentaţi grafic o realizare posibilă, pe parcursul a 10 paşi, generată de un amestec probabilistic de două distribuţii cu spaţiile de eşantionare $S_1 = \{-1;1\}$ şi $S_2 = \{0;2\}$.

* * *

Aplicaţia 11.3.7

Realizaţi un program care să genereze şi să salveze într-un fişier ASCII un eşantion de date conform unui model de amestec a 4 distribuţii i) exponenţiale cu medii diferite, ii) binomiale, cu număr identic de repetări, dar cu şanse diferite de succes.

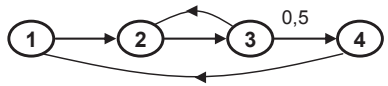
Programul se realizează parametrizat, utilizatorul putând fixa parametrii distribuţiilor, coeficienţii de amestec şi dimensiunea eşantionului.

* * *

11.3.3.2. Lanţuri Markov ascunse

Un alt exemplu de problemă complexă, ce poate fi rezolvată utilizând algoritmul EM, îl constituie estimarea parametrilor asociaţi unui **model cu lanţ Markov ascuns, HMM** (***H**idden **M**arkov **M**odel*). Un astfel de model presupune că, pentru generarea eşantionului de rezultate $\mathbf{y} = (y_1, \cdots, y_n, \cdots, y_N)$, unde n reprezintă momentul (pasul) în care apare rezultatul y_n, se utilizează K distribuţii distincte $\{\mathcal{D}_1, \mathcal{D}_2, \cdots, \mathcal{D}_K\}$, dintre care doar una conduce la apariţia

fiecărui rezultat. Alegerea distribuţiei se face pe baza indicilor acestora, precizaţi de stările în care se găseşte un lanţ Markov cu, desigur, K stări, în momentele de apariţie a rezultatelor.



Stare	Distribuţie
1	Normală (5, 10)
2	Binomială (1/2,10)
3	Geometrică (5/10)
4	Exponenţială (1/5)

Figura 11.3.1: Exemplu de model Markov ascuns

Aplicaţia 11.3.9

Alcătuiţi un posibil eşantion de 10 rezultate generate de modelul markovian ascuns, specificat în figura 11.3.1.

* * *

Aplicaţia 11.3.10

Realizaţi un program în MatLab care să genereze un eşantion de rezultate conform unui model cu lanţ Markov ascuns, în care procesul evoluează conform diagramei de tranziţii precizată în figura 11.3.2, iar stărilor li se asociază distribuţii Bernoulli de parametri q_1, respectiv q_2.

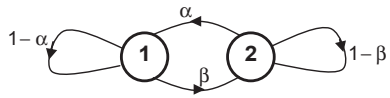


Figura 11.3.2: Diagrama de tranziţii a lanţului Markov din Aplicaţia 11.3.10

Programul se realizează parametric, atribuirea valorilor pentru mărimile caracteristice ($N, \alpha, \beta, q_1, q_2$) lăsându-se la latitudinea utilizatorului.

Indicaţii: se foloseşte funcţia MatLab `hmmgenerate()`, conform informaţiilor din fereastra de comandă, în urma comenzii `help hmmgenerate`, sau în pagina de referinţă din instrumentarul statistic (Statistics Toolbox), accesată în urma comenzii `doc hmmgenerate`.

* * *

11.3.3.3. Distribuţii matriceal-exponenţiale

O distribuţie cu faze, de genul celor prezentate în cadrul secţiunii 6.5, *"Metoda fazelor",* reprezintă, la modul general, distribuţia **timpului de absorbţie**, τ, ce caracterizează un lanţ Markov, $S(t)$, cu un număr de K **stări** (faze) **tranzitorii** (cu tranziţii de intrare şi de ieşire) şi o **stare absorbantă** (doar cu tranziţii de intrare). Pentru un astfel de lanţ, în care starea 0 se consideră starea absorbantă, matricea generatoare infinitezimală poate fi pusă sub forma:

$$\mathbf{Q}=\begin{pmatrix} 0 & \mathbf{0}_{1,K} \\ \boldsymbol{\gamma}_{K,1} & \boldsymbol{\Gamma}_{K,K} \end{pmatrix} \qquad (11.3.53)$$

în care blocurile partiţiei reprezintă:

- $\boldsymbol{\gamma}$ **= vectorul ieşirilor** (***exit vector***) ale cărui elemente, γ_i, sunt ratele de tranziţie din starea i, cu $i=\overline{1,K}$, în starea 0,
- $\boldsymbol{\Gamma}$ **= matricea generatoare de faze** ale cărei elemente, $\Gamma_{i,j}$, sunt ratele de tranziţii între stări tranzitorii, cu $i,j=\overline{1,K}$ şi $i\neq j$, iar $\Gamma_{i,i}$, cu $i=\overline{1,K}$, sunt ratele de ieşire din stările tranzitorii, cu semn schimbat.

Luând în considerare cele două mărimi, $\boldsymbol{\gamma}$ şi $\boldsymbol{\Gamma}$, matricea probabilităţilor de tranziţie de-a lungul unui interval, y, adică:

$$\mathbf{P}[y]=\exp(\mathbf{Q}\cdot y)=\sum_{i=0}^{\infty}\mathbf{Q}^i\cdot y^i/i! \qquad (11.3.54)$$

poate fi adusă la forma:

$$\mathbf{P}[y]=\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{1}-\exp(\boldsymbol{\Gamma}y)\cdot\mathbf{1} & \exp(\boldsymbol{\Gamma}y) \end{pmatrix} \qquad (11.3.55)$$

în care s-a ţinut cont că $\boldsymbol{\gamma}=-\boldsymbol{\Gamma}\cdot\mathbf{1}$, cu $\mathbf{1}^t=(1,1,\cdots,1)$, deoarece toate liniile din $\mathbf{Q}$ au suma nulă.

Prin urmare, ţinând cont şi de vectorul probabilităţilor iniţiale ale stărilor tranzitorii, $\mathbf{p}(0)$, se ajunge la următoarele caracteristici analitice ale timpului de absorbţie, adică ale unei distribuţii cu faze:

- **funcţia de distribuţie**:
$$F_\tau(y)=1-\mathbf{p}(0)\cdot\exp(\boldsymbol{\Gamma}\cdot y)\cdot\mathbf{1} \qquad (11.3.56)$$
- **densitatea de probabilitate**:
$$f_\tau(y)=\mathbf{p}(0)\cdot\exp(\boldsymbol{\Gamma}\cdot y)\cdot\boldsymbol{\gamma} \qquad (11.3.57)$$
- **transformata Laplace**:
$$\hat{\tau}(s)=\mathbf{p}(0)\cdot(s\cdot\mathbf{I}-\boldsymbol{\Gamma})^{-1}\cdot\boldsymbol{\gamma} \qquad (11.3.58)$$
- **momentul de ordinul *n***:
$$\mathrm{E}[\tau^n]=(-1)^n\cdot n!\cdot\mathbf{p}(0)\cdot\boldsymbol{\Gamma}^{-n}\cdot\mathbf{1} \qquad (11.3.59)$$

Relaţiile (11.3.56) – (11.3.59) arată că o distribuţie cu faze, denumită şi **distribuţie matriceal exponenţială** (***matrix exponential distribution***), urmare a faptului că o matrice este argument al exponenţialei, este complet specificată de mărimile $\mathbf{p}(0)$ şi $\boldsymbol{\Gamma}$, numărul parametrilor caracteristici fiind, în principiu, egal

cu $K(K+1)$. Concret, însă, în cazurile practice, majoritatea elementelor din $\mathbf{p}(0)$ şi $\boldsymbol{\Gamma}$, sunt nule, precum în exemplele de distribuţii cu 4 faze, din tabelul 11.3.2.

Tabelul 11.3.2: Conţinutul lui $\mathbf{p}(0)$ şi $\boldsymbol{\Gamma}$ în cazul unor distribuţii particulare cu faze

	$\mathbf{p}(0)$	$\boldsymbol{\Gamma}$
Hipo-exponenţială (m – Erlang)	$(1,0,0,0)$	$\begin{pmatrix} -\mu_1 & \mu_1 & 0 & 0 \\ 0 & -\mu_2 & \mu_2 & 0 \\ 0 & 0 & -\mu_3 & \mu_3 \\ 0 & 0 & 0 & -\mu_4 \end{pmatrix}$
Cox	$(\alpha_1=1,0,0,0)$	$\begin{pmatrix} -(\alpha_2+\beta_2)\mu_1 & \alpha_2\mu_1 & 0 & 0 \\ 0 & -(\alpha_3+\beta_3)\mu_2 & \alpha_3\mu_2 & 0 \\ 0 & 0 & -(\alpha_4+\beta_4)\mu_3 & \alpha_4\mu_3 \\ 0 & 0 & 0 & -\mu_4 \end{pmatrix}$
Hiper-exponenţială	$(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$	$\begin{pmatrix} -\mu_1 & 0 & 0 & 0 \\ 0 & -\mu_2 & 0 & 0 \\ 0 & 0 & -\mu_3 & 0 \\ 0 & 0 & 0 & -\mu_4 \end{pmatrix}$

Semnificaţiile notaţiilor folosite în tabel sunt:

- μ_i = rata de ieşire din starea i,
- α_i şi β_i = probabilităţile de selecţie a uneia din cele două tranziţii de ieşire din starea i.

În privinţa mediei timpului de absorbţie, a momentului de ordinul 1, aceasta se obţine calculând, mai întâi, vectorul $\mathbf{T}^t=(\tau_1,\tau_2,\cdots,\tau_K)^t=-\boldsymbol{\Gamma}^{-1}\cdot\mathbf{1}$ care conţine mediile timpilor de absorbţie, τ_i, când procesul pleacă din starea i. Urmează, apoi, înmulţirea cu vectorul probabilităţilor iniţiale de stare, $\mathbf{p}(0)$, adică aplicarea teoremei probabilităţilor totale, rezultatul final fiind media timpului de absorbţie, τ, indiferent de starea de plecare.

Aplicaţia 11.3.12

Precizaţi conţinutul vectorului de ieşire şi reprezentaţi grafic diagrama de tranziţii pentru fiecare distribuţie cu faze din tabelul 11.3.2.

* * *

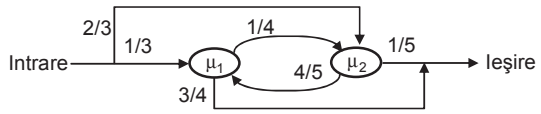


Figura 11.3.3: Exemplu de reţea în două faze

Aplicaţia 11.3.13

Se consideră reţeaua de faze din figura 11.3.3. Determinaţi funcţia de densitate de probabilitate, media şi varianţa variabilei aleatorie corespunzătoare, ştiind că $\mu_1 = \mu_2 = 4\sec^{-1}$. Ce valori au componentele vectorului $\mathbf{T}$?

Indicaţie: Calculul mediilor timpului de absorbţie, condiţionate de starea de plecare, se determină rezolvând sistemul $-\mathbf{\Gamma} \cdot \mathbf{T}^t = \mathbf{1}$. Ţinând cont de datele problemei, conţinutul matricei $\mathbf{\Gamma}$, este dat de produsul

$$\mathbf{\Gamma} = -\mathbf{M} \cdot (\mathbf{I} - \mathbf{P})$$

în care $\mathbf{M} = \mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_K)$ este **matricea ratelor de ieşire din faze**, iar $\mathbf{P} = \left\{ p_{i,j} \right\}$, cu $p_{i,i} = 0$ şi $i, j = \overline{1, K}$, este **matricea de rutare** în reţeaua de faze.

În urma substituirii şi prin rearanjare de termeni, se ajunge la următorul sistem: $\mathbf{T}^t = \mathbf{M}^{-1} \cdot \mathbf{1} + \mathbf{P} \cdot \mathbf{T}^t$, care motivează faptul că media timpului de absorbţie (de traversare a reţelei), începând din faza i, este egală suma dintre media fazei i şi mediile timpilor de absorbţie începând din stările la care se ajunge în urma tranziţiei, ponderate cu probabilităţile corespunzătoare, de rutare. Concret, în cazul reţelei analizate, una din ecuaţiile noului sistem este: $\tau_1 = 1/\mu_1 + 1/4 \cdot \tau_2$

Aplicaţia 11.3.19

Realizaţi un program care să înregistreze în format ASCII eşantioane de rezultate conform oricărui tip particular de distribuţii cu 4 faze.

* * *

11.3.3.4. Procese Markov de sosiri

Procesele Markov de sosiri **MAP** (***M**arkov **A**rrival **P**rocess*) au la bază, ca şi distribuţiile cu faze, un lanţ Markov continuu cu o stare absorbantă. Spre deosebire, însă, de acestea, în cazul proceselor Markov de sosiri, probabilităţile iniţiale de stare, de la un rezultat la altul, depind de ultima stare tranzitorie din care procesul a trecut prin starea absorbantă, cu ocazia generării rezultatului precedent.

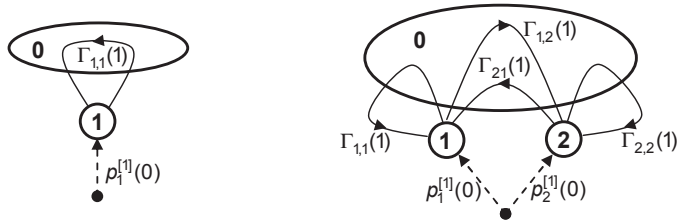
Modificarea specificaţiei, în sensul mai sus menţionat, constituie, în fapt, o generalizare a modelului de distribuţii cu faze, care a condus la o clasă de procese cu versatilitate sporită, ce pot modela fenomene complexe, cu rezultate corelate.

Caracterizarea unui proces Markov de sosiri se face prin intermediul următoarelor mărimi [Breuer] :

- $\mathbf{p}^{[1]}(0)$ - probabilităţile iniţiale ale stărilor (fazelor), care contribuie la generarea primului ([1]) rezultat,
- $\mathbf{\Gamma}(0)$ - matricea ratelor de tranziţie directă, între stările tranzitorii, $\Gamma_{i,j}$, cu $i \neq j$ şi a ratelor cu semn schimbat, de ieşire din stările tranzitorii, $\Gamma_{i,i}$,
- $\mathbf{\Gamma}(1)$ - matricea ratelor de tranziţie indirectă, între stările tranzitorii, care trec prin starea absorbantă, 0,

în care, argumentele matricelor $\mathbf{\Gamma}$ precizează, în fapt, dacă tranziţia este (când argumentul este 1) sau nu (când argumentul este 0) însoţită de o sosire.

Figura 11.3.6 prezintă, drept exemplificare, modul de reprezentare a diagramelor de tranziţii corespunzătoare procesului Poisson şi procesului Poisson modulat Markov, **MMPP** (***M**arkov **M**odulated **P**oisson **P**rocess*), văzute ca procese Markov de sosiri.



a) proces Poisson (PP) **b) proces Poisson modulat Markov (MMPP)**

Figura 11.3.6: Diagrame de tranziţii pentru 2 procese Markov particulare de sosire

În primul caz, există o unică stare tranzitorie şi o unică tranziţie care "străbate" starea absorbantă, ocazie cu care se generează o nouă sosire. În cazul al doilea, exemplul ia în considerare doar două stări tranzitorii între care nu există decât tranziţii prin starea absorbantă, ceea ce face ca procesul să genereze sosiri la fiecare tranziţie, ca şi în cazul anterior, dar cu rate care diferă funcţie de starea tranzitorie premergătoare.

Aplicaţia 11.3.21

Reprezentaţi grafic diagrama de tranziţii a unui proces Markov de sosiri cu două stări tranzitorii, care prezintă toate tranziţiile posibile.

Aplicaţia 11.3.22

Precizaţi în ce condiţii un proces Markov de sosiri, cu două stări, este identic cu modelul de distribuţie cu 2 faze, tip: i) hiper-exponenţial; ii) 2-Erlang.

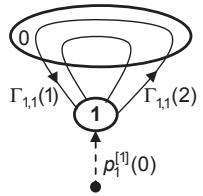
* * *

Aplicaţia 11.3.24

Realizaţi un program care să înregistreze într-un fişier cu extensia `txt`, rezultatele în format ASCII, generate de realizările particulare ale unui proces Markov de sosiri cu 4 stări tranzitorii.

11.3.3.5. Procese Markov de sosiri ponderate

Procesele Markov de sosiri ponderate **BMAP** (***B**atch **M**arkov **A**rrival **P**rocess*) constituie o extensie a proceselor Markov de sosiri, în sensul că o sosire este caracterizată, pe lângă momentul apariţie sale, şi de o anumită dimensiune. Mai exact, se consideră un set de ponderi, numere naturale între 1 şi M, iar fiecărei ponderi i se asociază câte o matrice de rate ale tranziţiilor ce străbat starea absorbantă.



a) diagramă de tranziţii

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} y_1 & y_2 & \cdots & y_N \\ z_1 & z_2 & \cdots & z_N \end{pmatrix}$$

$z_n = 1,2$ cu $n = \overline{1,N}$

b) eşantion particular de rezultate

Figura 11.3.7: Exemplu de proces Markov de sosiri ponderate

Figura 11.3.7 prezintă un exemplu de diagramă de tranziţii pentru un proces Markov de sosiri cu două ponderi, însoţită, alăturat, de modul de prezentare, în acest caz, a unui eşantion particular de rezultate. Valorile ponderilor sunt 1 şi 2, ceea ce înseamnă că fenomenul real, supus modelării, însoţeşte fiecare apariţie cu una din două valori posibile.

Procesele Markov de sosiri ponderate sunt caracterizate, conform celor precizate mai sus, de următorul set de parametrii:

- $\mathbf{p}^{[1]}(0) =$ probabilităţile iniţiale ale stărilor (fazelor), care contribuie la generarea primului rezultat ($[1]$),
- $\mathbf{\Gamma}(0) =$ matricea ratelor de tranziţie directă, între stările tranzitorii, $\Gamma_{i,j}$, cu $i \neq j$ şi a ratelor cu semn schimbat, de ieşire din stările tranzitorii, $\Gamma_{i,i}$,
- $\mathbf{\Gamma}(b) =$ matricea ratelor de tranziţie indirectă, între stările tranzitorii, care trec prin starea absorbantă, 0, conducând la o sosire de pondere b, cu $b = \overline{1,B}$.

Aplicaţia 11.3.28

Realizaţi un program care să înregistreze într-un fişier cu extensia `txt`, rezultatele în format ASCII, generate de realizările particulare ale unui proces Markov de sosiri cu 4 ponderi şi cu 4 stări tranzitorii.

11.4 Modelarea mobilităţii

Progresele spectaculoase, înregistrate în diversele domenii tehnico-ştiinţifice, aferente realizării echipamentelor de comutaţie şi transmisiuni, au permis ca, prin intermediul unui terminal specializat, **MS** (***M**obile **S**tation*), un abonat să aibă, în principiu, acces la reţeaua de telecomunicaţii indiferent de poziţia sa geografică, variabilă, în timp. Denumirea anterioară, ca şi celelalte noţiuni introduse în cele ce urmează, cu scopul descrierii reţelelor celulare, la nivel de principiu, sunt preluate din dicţionarul de termeni ai standardului **GSM**, (***G**lobal **S**ystem for **M**obile Communication*) [Ioan]. Această caracteristică, întâlnită desigur şi în cazul reţelelor radio mobile cu destinaţie specială (de exemplu anumite reţele militare sau ale serviciilor medicale de urgenţă), capătă un caracter *comun* în cazul ***reţelelor celulare*** în care o anumită *cale radio* (frecvenţă) *de acces* nu mai este rezervată unui anumit abonat, ci poate fi pusă la dispoziţia oricărui abonat care iniţiază la un moment dat un apel. Mai mult, prin adoptarea principiului ***reutilizării frecvenţelor***, ceea ce permite ca abonaţii aflaţi la distanţe corespunzătoare să utilizeze aceleaşi frecvenţe fără a se perturba reciproc, reţelele celulare conferă acestui gen de acces şi un ***caracter public***, de masă, numărul abonaţilor putând fi, în principiu, oricât de mare.

Reţelele celulare (figura 11.4.1) sunt alcătuite din punct de vedere topografic, conform denumirii, dintr-un ansamblu de ***celule*** (arii geografice de forme şi suprafeţe identice), care acoperă un anumit teritoriu şi realizează o "partiţie" a acestuia. Fiecare celulă este deservită de o ***staţie de bază***, **BS** (***B**ase **S**tation*), care asigură, pentru abonaţii aflaţi pe teritoriul său la un moment dat, accesul (emisie şi recepţie) la reţeaua de telecomunicaţii prin intermediul unui anumit număr de canale radio [ZoRS]. Canalele respective sunt repartizate atât cererilor de serviciu, iniţiate local, cât şi cererilor de ***handover*** (*handoff*) adresate celulei în cauză de surse de trafic deja în servire şi care, în urma deplasării lor, pătrund în aria de acoperire a acesteia.

Dat fiind faptul că benzile radio destinate acestui gen de acces cuprind un număr limitat de canale, repartizarea canalelor pe celule se face plecând de la principiul reutilizării frecvenţelor şi adoptând diverse soluţii (*politici*) de utilizare eficientă a acestora, precum:

- *împrumutul de canale* de la una din staţiile vecine, sau
- *alocarea flexibilă*, în care centrele de comutaţie ale reţelei mobile, MSC (***M**obile **S**witching **C**enter*) păstrează un anumit număr de canale (*guard channel*) pe care le distribuie temporar staţiilor de bază, funcţie de necesităţi.

Limitările tehnologice şi convenţionale ale spectrelor de frecvenţă, utilizate de reţelele de telecomunicaţii mobile, impun ca, indiferent de politica adoptată în distribuirea canalelor, o celulă să deţină un număr finit de canale, situaţie care conduce inevitabil la o respingere în proporţii, mai mult sau mai puţin neglijabile, a unei părţi din cererile de serviciu, adresate acesteia. În consecinţă, este justificat efortul de identificare sau elaborare de modele

matematice potrivite condiţiilor în care lucrează astfel de reţele şi care sunt necesare unei proiectări sau dezvoltări juste a acestora.

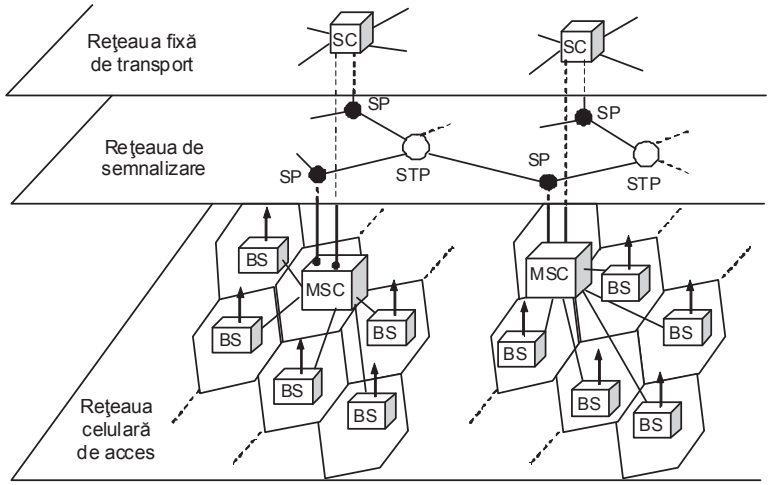


Figura 11.4.1: Integrarea reţelelor celulare în ansamblul reţelelor de telecomunicaţii
***SC** – centru de comutaţie (Switching Centre), **SP** – punct de semnalizare (Signalling Point), **STP** – punct de transfer de semnalizare (Signal Transfer Point), **MSC** – centru de comutaţie aparţinând reţelei celulare (Mobile Switching Centre), **BS** – staţie de bază (Base Station).*

Ipotezele fundamentale, luate în considerare în cazul evaluărilor de performanţă ale reţelelor fixe de acces se referă la natura procesului de sosire a cererilor de serviciu şi la duratele de servire a acestora şi conduc, de obicei, la aplicarea modelelor Erlang-B şi Engset (descrise în Capitolul 9). Aceste modele sunt dezvoltate în condiţii cât mai generale, care presupun că sosirile urmează un proces Poisson şi, în privinţa duratelor de servire, sunt interesate doar de valoarea lor medie [Doy89]. Prin urmare, valabilitatea lor se păstrează şi în cazul reţelelor celulare unde, însă, se ţine cont suplimentar şi de mobilitatea abonaţilor, care controlează caracteristicile modelelor, în speţă ***rata sosirii*** (rata cererilor de serviciu) şi ***durata medie a servirii***.

Stabilirea relaţiilor de calcul al celor două mărimi menţionate anterior pleacă de la considerarea, la modul general, a următoarelor variabile aleatorii:

- T_A - durata unui apel, din momentul iniţierii şi până în momentul finalizări,
- T_H - durata unui apel din momentul transferării şi până în momentul finalizării sale,

- T_{TA} - durata de tranzitare a unei celule de către o sursă de trafic (durata **de rezidenţă**), măsurată din momentul în care aceasta a iniţiat aici un apel,
- T_{TH} - durata de tranzitare a unei celule de către o sursă de trafic, măsurată din momentul în care aceasta a fost preluată de respectiva celulă, în urma unei cereri de handover.

Rata cererilor de serviciu adresate unei reţele celulare de telecomunicaţii se referă atât la iniţierea apelurilor, cât şi la solicitările de tip handover, corespunzătoare transferării apelurilor în curs de desfăşurare pe alte linii de acces (canale radio), funcţie de poziţia în timp a staţiilor mobile în cauză. Prin urmare, considerând omogene caracteristicile de trafic, de localizare şi de deplasare ale abonaţilor unei reţele celulare de telecomunicaţii, înseamnă că, la nivelul unei celule, se pot înregistra următoarele mărimi medii:

- λ_0 = rata cererilor de serviciu iniţiate local,
- λ_1 = rata cererilor de serviciu iniţiate local, dar care sunt transferate în primul rând înconjurător de celule, definită prin relaţia:

$$\lambda_1 = \lambda_0 \cdot P\left[T_A > T_{TA}\right] \qquad (11.4.1)$$

- λ_j = rata cererilor de serviciu iniţiate local, care ajung să fie transferate în cel de al *j*-lea rând înconjurător de celule, definită recursiv prin:

$$\lambda_j = \lambda_{j-1} \cdot P\left[T_H > T_{TH}\right] \qquad (11.4.2)$$

- λ_H = rata cererilor de handover adresate celulei, care, conform cu figura 11.4.2, ce evidenţiază ipotezele de omogenitate amintite anterior, are expresia:

$$\lambda_H = \lambda_1 + \lambda_2 + \ldots = \lambda_0 \cdot \frac{P\left[T_A > T_{TA}\right]}{1 - P\left[T_H > T_{TH}\right]} \qquad (11.4.3)$$

- λ_{tot} = - rata totală a cererilor de linii de acces (canale radio) adresate celulei, cu relaţia evidentă:

$$\lambda_{tot} = \lambda_0 + \lambda_H \qquad (11.4.4)$$

La modul general, în cazul în care variabilele implicate sunt considerate independente, cele două probabilităţi ce apar în relaţiile anterioare se pot exprima analitic conform teoremei probabilităţilor totale, sub forma:

$$P\left[T_* > T_{T_*}\right] = \int_{S_{T_{T_*}}} P\left[T_* > x\right] \cdot f_{T_{T_*}}(x)\,\mathrm{d}x \qquad (11.4.5)$$

unde $*$ poate fi A sau H, iar $S_{T_{T_*}}$ este spaţiul de eşantionare al variabilelor aleatorii specificate ca indice. Concret, relaţia anterioară se particularizează în funcţie de distribuţiile analitice sau empirice considerate.

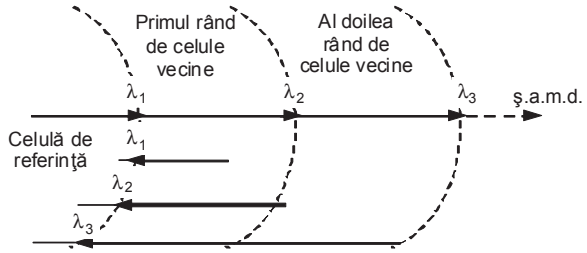


Figura 11.4.2: Omogenitatea teritorială evidenţiată prin echilibrul ratelor de handover

Astfel, în cazul adoptării celei mai familiare distribuţii, ce aproximează satisfăcător o paletă largă de aplicaţii şi oferă premisele unei analize matematice comode prin deţinerea proprietăţii de lipsă a memoriei, cele două variabile aleatorii generice, T_* şi T_{T_*}, sunt caracterizate de următoarele distribuţii de probabilitate:

$$f_{T_*}(x) = \mu_* \cdot \exp(-\mu_* x) \text{ şi } f_{T_{T_*}}(x) = \mu_{T_*} \cdot \exp(-\mu_{T_*} x) \qquad (11.4.6)$$

în care $1/\mu_*$ reprezintă durata medie de tranzit a unei celule, măsurată din momentul iniţierii apelului ($* \leftarrow A$), respectiv handover-ului ($* \leftarrow H$). Desigur, proprietatea sus menţionată conduce la concentrarea numărului de variabile aleatorii şi parametrii corespunzători, în sensul că:

$$\mu_A = \mu_H \overset{\text{notat}}{=} \mu \text{ şi } \mu_{TA} = \mu_{TH} \overset{\text{notat}}{=} \mu_T \qquad (11.4.7)$$

În consecinţă, revenind la relaţia (11.4.5) şi făcând substituţiile şi calculele de rigoare, rezultă că:

$$P[T > T_T] = \frac{\mu_T}{\mu + \mu_T} = \frac{1/\mu}{1/\mu_T + \dfrac{1/\mu}{1/\mu_T}} \qquad (11.4.8)$$

ultima exprimare punând în evidenţă un indicator al *dinamismului* surselor de trafic din punct de vedere al schimbării poziţiei în timp, numit ***mobilitate***, m:

$$m = \frac{1/\mu}{1/\mu_T} \qquad (11.4.9)$$

Reprezentarea grafică, din figura 11.4.3, a relaţiei (11.4.8) evidenţiază un lucru aşteptat şi anume că în reţelele cu mobilitate scăzută ($m \to 0$) sursele de trafic părăsesc rareori celula în care au iniţiat apelul înaintea finalizării acestuia, iar în cazul contrar, al mobilităţii ridicate ($m \to \infty$), de cele mai multe ori un apel

ajunge să fie transferat înaintea expirării sale. Funcţie de anumite praguri ale probabilităţii $P[T > T_T]$, precum cele considerate în figura 11.4.3, impuse mai mult sau mai puţin convenţional, se poate face o clasificare a mobilităţii surselor de trafic din reţelele celulare, de genul: *scăzută*, *moderată* sau *înaltă*.

Din punct de vedere practic, mobilitatea depinde în principal de cât de repede se mişcă sursele de trafic şi de cât de mici sunt dimensiunile celulelor.

Valoarea medie a duratelor de servire, pe parcursul cărora liniile de acces sunt angajate pentru a satisface diverse conexiuni, se poate determina experimental, prin măsurători repetate şi analize statistice corespunzătoare, sau analitic, prin identificarea unei variabile aleatorii, τ, a cărei distribuţie reprezentată prin $f_\tau(t)$ sau $F_\tau(t)$ se apropie satisfăcător de realitate şi aplicarea următoarei relaţii:

$$\mathrm{E}[\tau] = \int_{S_\tau} t \cdot f_\tau(t)\,\mathrm{d}t \qquad (11.4.10)$$

unde S_τ reprezintă spaţiul de eşantionare al variabilei aleatorii în cauză.

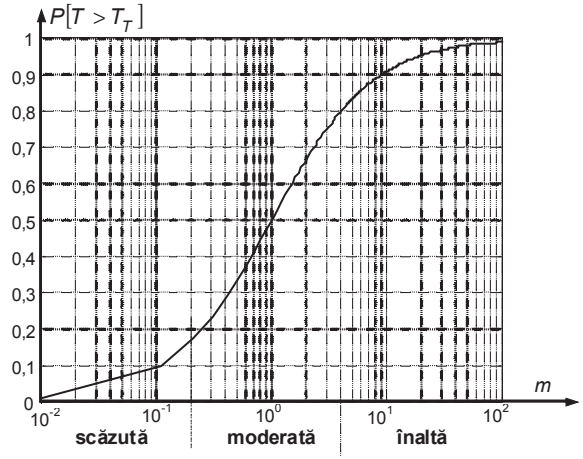


Figura 11.4.3: Variantă de clasificare a mobilităţii într-o reţea celulară

În cazul reţelelor fixe, distribuţia timpului de servire (angajare) a unei linii de acces este identică cu cea a duratelor pe parcursul cărora se desfăşoară apelurile, adică:

$$P[\tau \le t] = P[T_A \le t] \qquad (11.4.11)$$

Acest lucru nu mai este, însă, valabil în cazul reţelelor celulare, unde un apel poate "părăsi" o celulă înainte de momentul finalizării sale, eliberând linia de

acces ce i-a fost pusă la dispoziţie pe teritoriul celulei respective, apelul continuând să se deruleze, sub supravegherea celulei în care a fost transferat şi folosind resursele acesteia. Prin urmare, distribuţia timpului de servire este oferită de această dată de relaţia generală:

$$F_\tau(t) = p_A \cdot P\left[\min\{T_A, T_{TA}\} \le t\right] + p_H \cdot P\left[\min\{T_H, T_{TH}\} \le t\right] \quad (11.4.12)$$

unde p_A şi p_H sunt probabilităţile ca linia de acces (canalul) să fie distribuită unui apel, respectiv unui handover. Reprezentând, la modul empiric, raportul pe termen lung al unui anumit gen de cerere faţă de numărul total de cereri, cele două probabilităţi se definesc astfel:

$$p_A = \lambda_0/\lambda_t \text{ şi } p_H = \lambda_H/\lambda_t \quad (11.4.13)$$

Adoptând distribuţiile particulare considerate anterior, vezi relaţiile (11.4.6) şi (11.4.7), expresia (11.4.12) devine:

$$F_\tau(t) = P\left[\min(T, T_T) \le t\right] = 1 - P[T > t] \cdot P[T_T > t] = 1 - e^{-(\mu+\mu_t)\cdot t} \quad (11.4.14)$$

Prin urmare, media timpului de servire are, în acest caz, expresia:

$$E[\tau] = 1/(\mu + \mu_T) = 1/\mu \cdot 1/(1+m) \quad (11.4.15)$$

Figura 11.4.4 prezintă variaţia timpului mediu de servire raportat la durata medie a unui apel în funcţie de mobilitatea surselor de trafic. Cum era de aşteptat, într-o reţea imobilă ($m \to 0$), durata medie (şi nu numai) de servire (ocupare) a unei linii de acces corespunde cu cea a apelului, iar pe măsură ce mobilitatea creşte, durata medie de servire (referitoare la linia de acces) scade

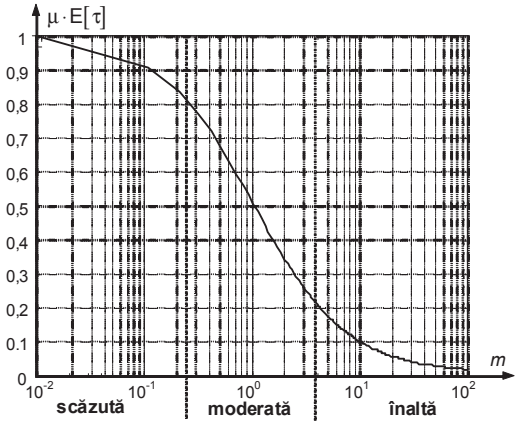


Figura 11.4.4: Media timpului de ocupare a liniilor de acces, normată cu durata medie de servire, dintr-o reţea celulară

din ce în ce mai mult, tinzând la limită spre zero. Pragurile indicate pe orizontală în figura 11.4.4 s-au stabilit plecând de la reperele de demarcare a domeniilor de mobilitate, ce au fost fixate în cadrul figurii 11.4.3.

Conform cu cele menţionate la începutul expunerii, privitoare la dimensionarea reţelelor celulare, indicatorii de performanţă necesari înfăptuirii acestui demers, au, în principiu, ca formule de calcul, expresiile dezvoltate în cadrul modelului Erlang-B. În consecinţă, diferenţa dintre planificarea unei reţele celulare faţă de planificarea unei reţele fixe nu constă în expresiile de calcul utilizate, ci în semnificaţia parametrilor λ şi $E[\tau]$, consideraţi, care semnifică:

- în cazul reţelelor celulare: rata totală de sosire a cererilor de apel şi de handover, respectiv durata medie de ocupare a unui canal radio,
- în cazul reţelelor fixe: rata de sosire a cererilor de apel, respectiv durata medie a unui apel.

De fapt, atât relaţia Erlang-B cât şi definiţia gradului de utilizare al serverelor (linii de acces) se referă, în cazul de faţă, al reţelelor celulare, doar la două mărimi: traficul oferit, ρ, şi numărul de servere. Deci, probabilitatea ocupării tuturor liniilor de acces ce aparţin unei celule depinde de valoarea traficului oferit, dat de expresia:

$$\rho = \lambda_t \cdot E[\tau] \quad (11.4.16)$$

care, particularizată în cazul ipotezelor asumate anterior, devine:

$$\rho = \underbrace{\lambda_0(1+m)}_{\lambda_t} \cdot \underbrace{\frac{1}{\mu} \cdot \frac{1}{1+m}}_{E[\tau]} = \frac{\lambda_0}{\mu} \quad (11.4.17)$$

Rezultatul este cumva surprinzător, dar totuşi de aşteptat întrucât traficul oferit în reţele celulare sau fixe, de acelaşi tip de surse, rămâne în principiu acelaşi, dacă nu ţinem seama de influenţa tarifelor practicate, o creştere a mobilităţii implicând durate mai mici de angajare a canalelor dar, în *compensaţie*, o repartizare mai frecventă a lor.

Aplicaţia 11.4.1

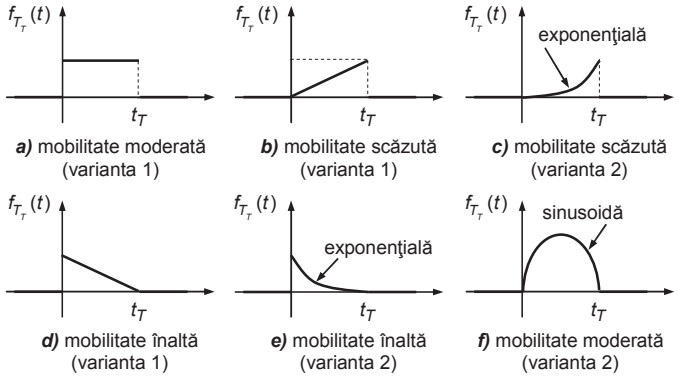
Să se determine formulele de calcul al caracteristicilor de trafic, propriu unei reţele celulare, dacă se respectă ipotezele precizate în prezentarea teoretică, cu excepţia celei referitoare la timpul de tranzit pentru care se ia în considerare o distribuţie a cărei funcţie de densitate de probabilitate este de tipurile reprezentate în cele 6 imagini din figura 11.4.5. Ce semnificaţie are mărimea t_T? Interpretaţi rezultatele pentru fiecare caz în parte.

Rezolvare: a) pentru densitatea de probabilitate cu aspectul propus, se obţine: $P[T_T \le t] = \begin{cases} t/t_T & \text{pentru } t \le t_T \\ 1 & \text{pentru } t > t_T \end{cases}$ şi

$$F_\tau(t) = P[\tau \le t] = \begin{cases} 1-(1-t/t_T)\,e^{-\mu t} & \text{pentru } t \le t_T \\ 1 & \text{pentru } t > t_T \end{cases},$$

adică: • $f_\tau(t) = \mu\, e^{-\mu t}\,(1 + 1/\mu' - \mu\, t/\mu')$, pentru $0 \le t \le t_T$ şi cu $\mu' = t_T \big/ \frac{1}{\mu}$,

• $\mathrm{E}[\tau] = \int_0^{t_T} (1 - F_\tau(t))\,dt = \int_0^{t_T} (1 - t/t_T)\, e^{-\mu t}\,dt = \frac{1}{\mu}\left[1 - 1/\mu' \cdot \left(1 - e^{-\mu'}\right)\right]$



a) mobilitate moderată (varianta 1) ***b)*** mobilitate scăzută (varianta 1) ***c)*** mobilitate scăzută (varianta 2)

d) mobilitate înaltă (varianta 1) ***e)*** mobilitate înaltă (varianta 2) ***f)*** mobilitate moderată (varianta 2)

Figura 11.4.5: Reţea celulară cu diferite grade de mobilitate

11.5 Aplicaţii suplimentare

Aplicaţia 11.5.1

Determinaţi rata medie a unui proces IPP, în care lanţul Markov modulator urmează diagrama din figura 11.2.2.3.

Caz particular $a = 0,6/\text{sec}$, $s = 0,2/\text{sec}$ şi $\lambda = 10$ pachete/sec.

Aplicaţia 11.5.2

O sursă de trafic poate genera un pachet la intervale succesive de timp egale cu $1\ \mu\text{sec}$. Considerând că probabilitatea de generare a pachetului poate avea valorile $1/5$, $2/5$ sau $4/5$ în funcţie de starea în care se află la un moment dat un proces de creştere/descreştere cu trei stări şi cu tranziţii ascendente de rată $\alpha = 0,01\ \mu\text{sec}^{-1}$ şi tranziţii descendente de rată $\varpi = 0,02\ \mu\text{sec}^{-1}$, să se determine: debitul mediu, debitul de vârf şi coeficientul de rafală al sursei.

Aplicaţia 11.5.3

Fie o sursă a cărui debit poate varia conform diagramei din figura 11.5.1.

Să se determine valorile parametrilor care o caracterizează (r_p, r_m etc.).

Se dau: $\lambda = 20$ celule/msec, $\varpi = 25\ \text{msec}^{-1}$ şi $\alpha = 10\ \text{msec}^{-1}$.

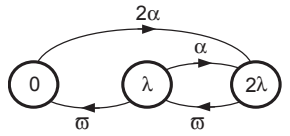


Figura 11.5.1: Caz particular de proces MMPP

Aplicaţia 11.5.4

Debitul unei surse de trafic VBR se modifică conform unui proces Poisson, a cărui rată poate lua valorile 2, 4 sau 8 Mbiţi/sec în funcţie de starea unui proces de creştere/descreştere, cu trei stări şi cu tranziţii ascendente de rată $\alpha = 0,5\,\text{sec}^{-1}$ şi tranziţii descendente de rată $\varpi = 0,4\,\text{sec}^{-1}$.

Determinaţi debitul mediu, debitul de vârf şi coeficientul de rafală ce caracterizează sursa de trafic în discuţie.

Aplicaţia 11.5.5

Determinaţi coeficientul de rafală ce caracterizează o sursă de trafic de tip ON/OFF: a) vocal (telefonic); b) de date.

Aplicaţia 11.5.6

În urma observării, pe durata a 4 ore, a unui grup de 4 unităţi de servire, s-au trasat diagramele corespunzătoare, de utilizare, prezentate în figura 11.5.2, în care simbolurile 1 şi 0 reprezintă stările de ocupare, respectiv neutilizare a acestora.

Stabiliţi caracteristicile traficului servit de respectivul grup pe parcursul intervalului de timp considerat.

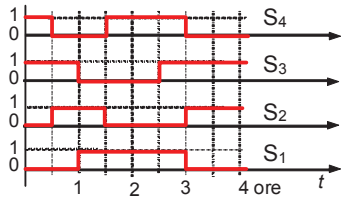


Figura 11.5.2: Diagramă de utilizare pentru 4 unităţi de servire

capitolul 12

Sisteme de servire cu aşteptare

Ori de câte ori ne găsim în ipostaza de ***client***, aceasta se datorează faptului că dorim să ni se ofere un ***serviciu***. Din nefericire, însă, în majoritatea cazurilor servirea propriu-zisă, cu toate eventualele sale neajunsuri, este condiţionată de un fenomen deseori mult mai greu de suportat şi anume ***aşteptarea***. Aşteptarea la secţia financiară pentru plata impozitelor (unde aşa zisul serviciu constă în primirea sumei în cauză de către funcţionarul din spatele ghişeului în schimbul unei chitanţe fiscale), aşteptarea autobuzului în care să te poţi urca pentru a ajunge la timp la serviciu (dacă nu cumva, pe parcurs, ai neşansa unor congestionări de trafic care îngreunează circulaţia rutieră alterând uneori insuportabil calitatea servirii), aşteptarea tonului de disc la miez de noapte cu prilejul anului nou, aşteptarea transferului unui fişier printr-o reţea suprasolicită în orele de vârf, şi aşa mai departe căci exemplele pot continua la nesfârşit, luând în considerare orice domeniu ale cărui resurse sunt limitate în timp şi spaţiu.

În toate aceste situaţii, ***clienţii*** nu sunt mulţumiţi că trebuie să aştepte, iar operatorii serviciilor respective nu sunt încântaţi că au şiruri de clienţi, deoarece pentru asigurarea unei fluenţe acceptabile trebuie să cheltuiască suplimentar pentru sporirea facilităţilor de servire. Desigur, termenul ***client*** este folosit în general şi el nu implică neapărat o fiinţă umană. Din exemplele anterioare se desprinde clar ideea că un client poate fi un om, un program de calculator, un apel telefonic, o informaţie (vocală sau codificată electric) etc.

Din aceste motive, luând în considerare interesele (cerinţele) clienţilor privind servirea, **QoS** (***Q**uality **o**f **S**ervice*), şi pe cele ale furnizorilor de servicii legate, printre altele, de utilizarea cât mai eficientă a resurselor disponibile, echipele de proiectare şi cercetătorii ştiinţifici din domeniul reţelelor de telecomunicaţii au conceput diverse tehnici de servire care să asigure un compromis acceptabil din punctul de vedere al ambelor părţi. Dintre acestea, cele mai larg folosite sunt ***concentrarea traficului***, ***multiplexarea deterministă în timp şi frecvenţă şi multiplexarea statistică***, tehnici care pot fi utilizate individual sau combinat (cazul tehnicii **CDMA** (***C**ode **D**ivision **M**ultiple **A**cces*) [Băn99], aplicată în reţelele mobile, care *concentrează* într-o bandă de frecvenţă, prin multiplexare în frecvenţă, un anumit număr de căi de convorbire).

Pentru analiza, proiectarea şi optimizarea într-un mod coerent a sistemelor de telecomunicaţii bazate pe aceste principii, s-a dezvoltat o teorie de sine-stătătoare, intitulată ***teoria aşteptării*** sau ***teoria şirurilor***, ce conţine un ansamblu de modele probabilistice. Aceste modele [Jew95], [Klein78], [Mih78] s-au elaborat plecând de la o serie de ipoteze privind sosirea clienţilor, desfăşurarea servirii şi echiparea sistemului şi oferă formule de calcul pentru diverşi indicatori de performanţă precum: număr mediu de clienţi în sistem, timpul mediu petrecut de un client în sistem sau probabilitatea ocupării complete a sistemului.

Matematicianul danez A. K. Erlang este considerat pionierul acestei teorii, studiind în mod special congestia traficului telefonic ("*The Theory of Probabilities and Telephone Conversations*" – 1909). Activitatea lui a fost continuată prin contribuţiile importante aduse de numeroşi teoreticieni din întreaga lume: E.C. Mollina (1927), A.N. Kolmogorov (1931), C.D. Crommelin (1932), F. Pollaczek (1932), A.Y. Khintchine (1932), D.G. Kendall (1951), J.R. Jackson (1954), P.J. Burke (1956), R.I. Wilkinson (1956), C. Palm (1957), R. Syski (1960), J.D.G. Little (1961), P. Le Gall (1962), Gh. Mihoc (1962), V.E. Beneš (1963), A.M. Lee (1966), F.P. Kelly (1976), J.W. Cohen (1982), C.M. Harris (1998) etc.

În general, previziunile oferite de modelele probabilistice, precum şi ipotezele ce le condiţionează, sunt verificate practic în majoritatea situaţiilor şi, în plus, pot fi utilizate diverse mijloace de menţinere a comportamentului real al clienţilor în limitele modelului; în acest sens poate fi amintită, de exemplu, dotarea sistemelor de comunicaţii cu servicii de genul poştă electronică, poştă vocală, răspuns automat, precum şi cu posibilitatea de modificare orară a tarifelor. Există însă situaţii în care, deşi au fost luate măsuri preventive, anumiţi factori perturbatori modifică uneori comportamentul clienţilor, îndepărtându-l de modelul probabilistic adoptat, ajungându-se astfel la situaţii neprevăzute de acesta. De exemplu, *blocarea* unui sistem de servire poate fi cauzată de activitatea neaşteptată a clienţilor, în orele de vârf, mult peste activitatea "medie", considerată la proiectare, sau de anumite evenimente sociale, care au ca efect repetarea condiţionată şi perseverentă a cererilor nesatisfăcute (fenomen pe care modelul ce a stat la baza dimensionării respectivului sistem nu l-a considerat).

12.1 Elementele de modelare

În vederea modelării sistemelor de servire cu aşteptare, s-a căutat identificarea unei structuri cât mai abstracte care să prezinte caracteristicile esenţiale ale oricărei situaţii concrete, întâlnită în realitate.

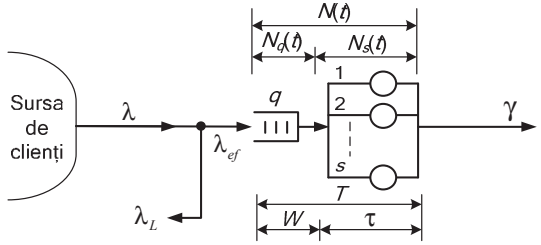


Figura 12.1.1: Elementele de modelare (structurale şi dinamice)

Astfel, conform figurii 12.1.1, se consideră că un *sistem de servire cu aşteptare* este *alimentat* de o *sursă* de clienţi, sursă cu caracteristici mai mult sau mai puţin dependente de sistemul de prelucrare, şi dispune de una sau de mai multe *unităţi de servire* (*servere*), identice din punct de vedere al modului de funcţionare (preluare clienţi, servire clienţi etc.). În plus, sistemul permite poziţionarea într-un şir de aşteptare a clienţilor sosiţi în timp ce toate serverele sunt ocupate cu servirea clienţilor aflaţi deja în sistem. Clienţii din şir vor fi preluaţi pentru a fi serviţi pe măsura disponibilizării serverelor şi după o anumită disciplină. Pentru şirul de aşteptare se poate impune o lungime limită maximă, ceea ce conduce la apariţia unor situaţii de blocare a sistemului, când o serie de clienţi nu vor putea fi primiţi în sistem (toate serverele sunt ocupate şi şirul este complet încărcat). Reprezentând componentele structurale de bază, unităţile de servire şi poziţiile (locaţiile) de aşteptare constituie împreună ***resursele*** de care dispune un anumit sistem de servire cu aşteptare.

În afară de elementele de configurare, modelarea sistemelor de servire cu aşteptare ia în considerare şi o serie de aspecte (elemente) dinamice, menţionate de asemenea în reprezentarea grafică din figura 12.1.1 şi având următoarele semnificaţii:

- λ = rata de sosire a clienţilor la intrarea în sistem [clienţi/sec] (rata cu care sursa externă oferă clienţi);
- λ_L = rata de respingere (de pierdere – ***L**oss*) a clienţilor;
- λ_{ef} = rata ***efectivă*** de intrare a clienţilor în sistem;
- γ = rata de plecare a clienţilor ce au fost acceptaţi şi serviţi în sistem, numită şi ***productivitate*** a sistemului. Desigur că $\gamma = \lambda_{ef}$, toţi clienţii intraţi în sistem, părăsindu-l mai devreme sau mai târziu;
- $N_q(t)$ = procesul aleatoriu corespunzător numărului de clienţi aflaţi în şirul de aşteptare (***q**ueue*), de lungime maximă q;

- $N_s(t)$ = procesul aleatoriu corespunzător numărului de clienţi aflaţi în ***s**ervire*, în cele s servere ale sistemului;
- $N(t)$ = procesul aleatoriu corespunzător numărului total de clienţi prezenţi în sistem (în serviciu sau în aşteptare);
- W = variabila aleatorie care reprezintă timpul de aşteptare în şir (***w**aiting time*);
- τ = variabila aleatorie care reprezintă timpul propriu-zis de servire (*service time*);
- T = variabila aleatorie care reprezintă timpul total petrecut de un client în sistem, numit şi timpul de întârziere în sistem (*delay time*) sau timpul de tranzit prin sistem sau de staţionare în sistem. În plus, modelul este specificat şi de *disciplina cozii*, care are o influenţă importantă asupra mărimilor W şi T. Indiferent însă care este această disciplină, între diferitele mărimi temporale, pentru orice client i, există evident relaţia:

$$T_i = W_i + \tau_i \qquad (12.1)$$

Performanţa sistemului este exprimată prin diferiţi indicatori, care privesc:

❖ ***punctul de vedere al utilizatorului***, precum:

- caracteristicile statistice (distribuţia, media, varianţa) ale timpilor W şi T, care asigură mai mult sau mai puţin calitatea dorită de satisfacere a serviciului (de exemplu, transferul unui fişier prin reţea). Desigur că, în cazul apelurilor telefonice, mărimea T este în principal determinată de participanţii la convorbire;
- raportul λ_L/λ, care este o măsură a neşansei ca cererea unui client să fie respinsă de un sistem.

❖ ***punctul de vedere al furnizorului de servicii***, precum:

- traficul servit sau gradul de utilizare al fiecărei unităţi de servire (empiric definit ca raport pe termen lung între durata ocupării şi durata totală de exploatare, fiind la limită probabilitatea de ocupare a acesteia), ce constituie o măsură a eficienţei în utilizarea lor de către sistem;
- rata efectivă de intrare a clienţilor în *sistem* $\lambda_{ef} = \lambda - \lambda_L$;
- caracteristicile statistice ale mărimilor $N(t)$, $N_q(t)$ şi $N_s(t)$, ce sunt utile, de exemplu, în dimensionarea sistemului.

12.2 Formula lui Little

Una dintre cele mai simple şi extrem de utilă relaţie, folosită frecvent în calculul indicatorilor de performanţă a sistemelor cu aşteptare este ***formula lui Little***, care se enunţă astfel: pentru un sistem ajuns la echilibru (procesele asociate sunt staţionare în sens larg), media numărului de clienţi în sistem, $\mathrm{E}[N]$, este produsul dintre rata de intrare în sistem, λ_{ef}, şi media timpului petrecut în sistem, $\mathrm{E}[T]$:

$$\mathrm{E}[N] = \lambda_{ef} \cdot \mathrm{E}[T] \qquad (12.2.1)$$

Demonstraţia acestei formule se poate face simplu, folosind schema generală a servirii din figura 12.2.1 în care:

- $N_a(t)$ = numărul clienţilor intraţi (sosiţi) în sistem până la momentul t;
- $N_d(t)$ = numărul clienţilor plecaţi până la momentul t;
- $N(t)$ = numărul clienţilor în sistem la momentul t.

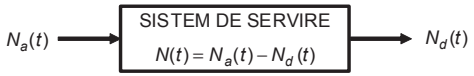


Figura 12.2.1: Schema generală a servirii

Fie că un client oarecare i intră în sistem la momentul $T_i^{(a)}$ şi părăseşte sistemul la momentul $T_i^{(d)}$, atunci timpul petrecut de acesta în sistem este:

$$T_i = T_i^{(d)} - T_i^{(a)} \qquad (12.2.2)$$

Presupunând că la momentul iniţial $t_0 = 0$ sistemul este gol, înseamnă că numărul clienţilor la momentul t este:

$$N(t) = N_a(t) - N_d(t) \qquad (12.2.3)$$

Fără a restrânge generalitatea demonstraţiei, se consideră o disciplină FIFO de servire a clienţilor, adică este posibilă o realizare pentru procesul $N_a(t)$ şi respectiv $N_d(t)$ ca în figura 12.2.2. Media în timp a numărului de clienţi din sistem este dată de relaţia:

$$\langle N \rangle_t = \frac{1}{t}\int_0^t N(t')\mathrm{d}t' = \frac{1}{t}\sum_{i=1}^{N_a(t)} T_i = \frac{1}{t}\cdot\left[\sum_{i=1}^{N_a(t)} T_i^{(d)} - \sum_{i=1}^{N_a(t)} T_i^{(a)}\right] \qquad (12.2.4)$$

care pune în evidenţă următoarele aspecte:

- restricţia $N(t_0) = 0$ nu este necesară;
- tipul disciplinei nu este important;
- numărul resurselor nu intervine în formulă.

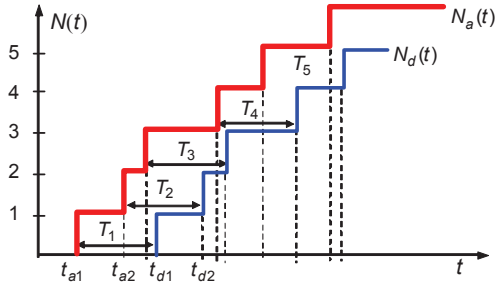


Figura 12.2.2: Procese de sosire şi de plecare a clienţilor

Dar rata sosirilor în intervalul $[0,t]$ este $\langle \lambda_{ef} \rangle_t = N_a(t)/t$, iar media timpului petrecut în sistem de $N_a(t)$ clienţi este:

$$\langle T \rangle_t = \frac{1}{N_a(t)} \cdot \sum_{i=1}^{N_a(t)} T_i \qquad (12.2.5)$$

ceea ce înseamnă că:

$$\langle N \rangle_t = \langle \lambda_{ef} \rangle_t \cdot \langle T \rangle_t \qquad (12.2.6)$$

Presupunând că toate procesele ataşate sistemului (sosiri de clienţi, timpi de staţionare) sunt ergodice, expresia (12.2.6) capătă de fapt forma $\mathrm{E}[N] = \lambda_{ef} \cdot \mathrm{E}[T]$ a formulei lui Little. Trebuie precizat că formula este aplicabilă în condiţii foarte generale pentru diferite sisteme de prelucrare, deoarece lipsesc din deducerea ei orice specificaţii suplimentare ergodicităţii.

Aplicaţia 12.2.1

Fie reţeaua de sisteme cu aşteptare *fără pierderi* din figura 12.2.3. Considerând că duratele medii de întârziere în fiecare sistem sunt cunoscute şi că în reţea se găseşte un singur client, să se determine numărul mediu de clienţi din fiecare sistem. Caz particular: $\mathrm{E}[T_1] = 4\,\mathrm{sec}$; $\mathrm{E}[T_2] = 6\,\mathrm{sec}$; $\mathrm{E}[T_3] = 10\,\mathrm{sec}$.

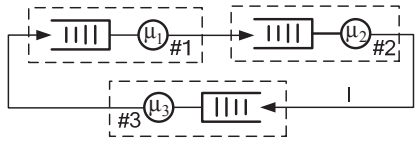


Figura 12.2.3: Reţea închisă cu 3 sisteme de servire cu aşteptare

Rezolvare: Durata medie a unui ciclu este $\mathrm{E}[T] = \sum_{i=1}^{3} \mathrm{E}[T_i] = 20\,\mathrm{sec}$, deci formula Little oferă, la nivel de reţea: $\lambda = \mathrm{E}[N]/\mathrm{E}[T] = 0{,}05\,\text{clienţi/sec}$. Deoarece reţeaua este presupusă fără pierderi, cu o structură închisă şi fără întoarceri, această rată se regăseşte la intrarea fiecărui sistem, determinând un număr mediu de clienţi pe sisteme $\mathrm{E}[N_i] = \lambda \cdot \mathrm{E}[T_i]$ şi pe ansamblul reţelei: $\sum_{i=1}^{3} \mathrm{E}[N_i] = 1\,\text{client}$.

* * *

Această aplicaţie pune în evidenţă faptul că, din punct de vedere al formulei Little, orice subsistem al unui sistem poate fi "tratat" ca un sistem de sine stătător. Astfel, referindu-ne la figura 12.1.1, se pot considera componentele modelului ca subsisteme distincte ale unui ansamblu global. În consecinţă, formula lui Little se particularizează funcţie de variabilele strict ataşate, fie şirului de aşteptare, fie serverelor, conducând la următoarele rezultate:

- *numărul mediu de clienţi aflaţi în şirul de aşteptare* este egal cu produsul

dintre rata medie de acceptare în sistem, λ_{ef} (*rata de intrare efectivă*) şi timpul mediu de aşteptare:

$$\mathrm{E}\left[N_q\right] = \lambda_{ef} \cdot \mathrm{E}[W] \qquad (12.2.7)$$

- *numărul mediu de clienţi aflaţi în serviciu* este egal cu produsul dintre rata medie de acceptare a clienţilor în sistem şi timpul mediu de servire a acestora:

$$\mathrm{E}[N_s] = \lambda_{ef} \cdot \mathrm{E}[\tau] \qquad (12.2.8)$$

Aplicaţia 12.2.2

Fie un sistem de servire care dispune de un singur server. Determinaţi expresia de calcul al traficului servit de acesta (gradul său de utilizare) funcţie de rata de tranzit a clienţilor prin sistem şi de durata medie de servire a acestora.
Rezolvare: Gradul de utilizare a unui server este probabilitatea sa de ocupare, fiind echivalent în cazul acestei aplicaţii cu probabilitatea $P[N_s(t) = 1]$.

12.3 Clasificarea Kendall

Pentru referirea într-un mod clar, simplu şi precis la un anumit model probabilistic caracteristic unui anumit gen de sistem de servire cu aşteptare, se foloseşte o codificare consacrată, ***clasificarea Kendall***, ce presupune ataşarea unei etichete cu cel mult 6 caractere alfa-numerice fiecărui tip de model. Procedura iniţială de clasificare este datorată lui D. G. Kendall, care a propus drept criterii de clasificare procesul de sosire, repartiţia timpului de servire şi numărul unităţilor de servire, observând că acestea sunt elementele principale ce particularizează un anumit model.

În prezent, cele 6 caractere ale clasificării Kendall au semnificaţiile prezentate în schema din figura 12.3.1, după cum urmează:

1. *Procesul de sosire* este specificat prin câteva notaţii consacrate şi anume: **M** - lege exponenţială; **D** - lege deterministă; $\mathbf{E_k}$ - lege Erlang de ordin k; $\mathbf{H_n}$ - lege hiperexponenţială de ordin n; **G** - lege generală.

De obicei, modelarea sosirilor întâmplătoare ia în considerare, pentru a stabili ***timpul*** θ ***dintre două sosiri succesive***, legea exponenţială:

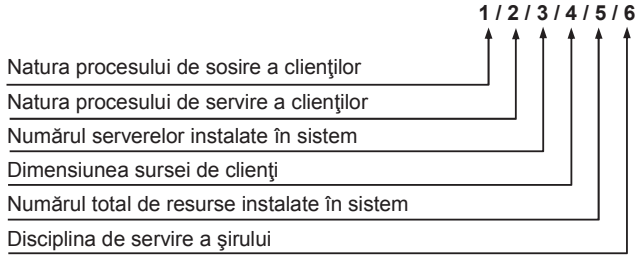

1 / 2 / 3 / 4 / 5 / 6
Natura procesului de sosire a clienţilor
Natura procesului de servire a clienţilor
Numărul serverelor instalate în sistem
Dimensiunea sursei de clienţi
Numărul total de resurse instalate în sistem
Disciplina de servire a şirului

Figura 12.3.1: Simbolurile clasificării Kendall

$$P[\theta \le t] = 1 - \exp(-\lambda \cdot t) \qquad (12.3.1)$$

caz în care sosirea, văzută ca proces de numărare pe parcursul unui interval t este modelată de o distribuţie Poisson de acelaşi parametru λ:

$$p_n(t) = \frac{(\lambda \cdot t)^n}{n!} \mathrm{e}^{-\lambda \cdot t} \qquad (12.3.2)$$

caz în care timpul mediu inter-apeluri este $\mathrm{E}[\theta] = 1/\lambda$.

2. *Legea de servire* se referă la distribuţia timpilor τ de servire a clienţilor admişi în sistem şi, pentru diferite moduri de servire, este specificat prin aceleaşi notaţii ca şi procesele de sosire. În multe aplicaţii funcţia de densitate de probabilitate pentru timpul de servire se consideră exponenţială:

$$f_\tau(t) = \mu \ \mathrm{e}^{-\mu \cdot t}, \ t \ge 0 \qquad (12.3.3)$$

caz în care timpul mediu de servire este $\mathrm{E}[\tau] = 1/\mu$.

3. *Numărul serverelor instalate în sistem*. Raportat la numărul s de servere prevăzute, sistemele de servire se împart în:

- *sisteme cu un singur server*, $s = 1$ (sisteme uni-server) – ceea ce înseamnă că un singur client poate fi servit la un moment dat,

- *sisteme cu mai multe servere* (sisteme multi-server) – care au $s > 1$ servere, toate cu caracteristici identice de servire şi care sunt alimentate echiprobabil din şirul comun de aşteptare; înseamnă deci că în acelaşi timp pot fi serviţi cel mult s clienţi.

4. *Dimensiunea sursei de clienţi (numărul maxim al clienţilor potenţiali oferiţi sistemului)* – este dată de numărul c al clienţilor potenţiali, în principiu cu comportament independent unul faţă de celălalt, conţinuţi în sursa corespunzătoare. Comparând dimensiunea sursei cu aceea a grupului de servere din sistemul de servire, sursele pot fi de două categorii:

- *sursă infinită*, dacă $c > 10 \cdot s$. Înseamnă că şirul de aşteptare poate deveni deosebit de lung, teoretic infinit, dacă sistemul de servire acceptă acest lucru.

- *sursă finită*, dacă $c \le 10 \cdot s$. În acest caz, numărul maxim al clienţilor prezenţi în sistem este totdeauna limitat la dimensiunea sursei.

5. *Numărul total de resurse* – este determinat de dimensiunea q a şirului de aşteptare adunată cu dimensiunea s a grupului de servere, adică $r = q + s$ (suma precizează de fapt numărul maxim de clienţi ce pot fi prezenţi în sistem la un moment dat).

În privinţa *şirului de aşteptare*, acesta poate fi de:

- *capacitate infinită* ($q = \infty$): orice client care soseşte are permisiunea să intre în sistemul de prelucrare, indiferent câţi alţii asemenea lui se află aici (**sisteme ideale, cu aşteptare şi menţinere**).

- *capacitate zero* ($q = 0$): clientul care soseşte atunci când toate serverele sunt ocupate nu este admis în sistem, devenind un client pierdut. Este cazul **sistemelor cu pierderi nete**.

- *capacitate finită* ($0 < q < \infty$): clientul care soseşte într-un moment de blocare a grupului de servere este primit în sistem şi aşteaptă servirea doar în cazul în care mai este vreo poziţie liberă în şir, altfel el este respins de sistem. Este cazul **sistemelor reale, cu aşteptare şi pierderi**.

6. *Disciplina de servire a şirului de aşteptare* – stabileşte regula de alegerea a clienţilor din şir pentru a fi serviţi. Se cunosc diferite asemenea discipline, dintre care cele mai frecvente sunt următoarele:

- **FIFO** (***F**irst **I**n **F**irst **O**ut*) sau **FCFS** (***F**irst **C**ome, **F**irst **S**erved*)
- **LIFO** (***L**ast **I**n **F**irst **O**ut*) sau **LCFS** (***L**ast **C**ome, **F**irst **S**erved*)
- **RSS** (***R**andom **S**election for **S**ervice*),
- **PR** (***P**riority*),
- **G** (***G**eneral*).

Observaţie: Deseori în practică, multe dintre modelele uzuale sunt reprezentate printr-un număr redus de simboluri ale clasificării Kendall. Acest lucru este permis deoarece simbolurile omise au următoarele valori implicite:

$$q = \infty,\ c = \infty,\ r = s + q = \infty,\ \text{FIFO}.$$

Aplicaţie 12.3.1

Precizaţi care sunt caracteristicile următoarelor modele de sisteme de servire cu aşteptare: M/M/3, M/D/1/8, M/G/2, D/M/1, G/D/1, D/D/1, $M/H_2/1$.

12.4 Modelul M/M/1

Modelul M/M/1 din punct de vedere teoretic poate fi considerat ca o particularizare a modelului M/M/s analizat anterior, deoarece el se referă la cazul special al unui sistem uniserver. Din punct de vedere practic el este deosebit de convenabil pentru studierea şi evaluarea performanţelor reţelelor de telecomunicaţii, în care nodurile au pe porturile de acces echipamente specializate emisie/recepţie, funcţionând ca un server dotat cu o memorie proprie. Prezentarea modelului se face stabilind iniţial, prin intermediul aplicaţiei ce urmează, contextul teoretic în care modelul M/M/1 se aplică.

Aplicaţia 12.4.2

Să se caracterizeze modelul M/M/1 din punct de vedere al terminologiei Kendall**.**

Rezolvare: Acest model consideră că:

1. sosirea clienţilor urmează un proces Poisson, de exemplu de rată medie λ, ceea ce înseamnă că: timpii dintre două sosiri succesive sunt variabile aleatorii independente şi identic distribuite (*i.i.d.*), care urmează o distribuţie exponenţială de medie $1/\lambda$;
2. timpii de servire sunt variabile aleatorii independente identic distribuite, ce urmează o distribuţie exponenţială, de medie $1/\mu$;

3. numărul de servere din sistem este egal cu $s = 1$;
4. sursa care alimentează sistemul este infinit de mare, $c = \infty$;
5. numărul clienţilor ce pot fi acceptaţi în sistem este infinit, $r = q + s = \infty$;
6. disciplina servirii este FIFO.

* * *

În condiţiile sus menţionate, procesul $N(t)$ *număr de clienţi în sistem* este un lanţ Markov continuu în timp. Într-adevăr, proprietatea de *lipsă a memoriei* proprie variabilelor aleatorii exponenţiale, ce caracterizează atât sosirile cât şi plecările din sistem, face ca timpul până la următoarea sosire, respectiv până la următoarea plecare, să fie independent de momentul de referinţă, deci procesul să depindă doar de prezent.

În plus, el face parte din categoria specială de procese Markov, şi anume aceea a proceselor de naştere şi moarte. Ultima afirmaţie este demonstrată pe baza următoarelor rezultate, ce iau în considerare numărul $N_a(t)$ de sosiri în intervalul t, numărul $N_d(t)$ de plecări în intervalul t şi un intervalul de timp δ foarte mic:

- probabilitatea ca doar un client să sosească în intervalul δ este semnificativă, depinzând doar de rata de sosire şi de lungimea intervalului:

$$P\left[N_a(\delta) = 1\right] = \frac{\lambda \cdot \delta}{1!} e^{-\lambda \cdot \delta} = \lambda \cdot \delta \cdot (1 - \lambda \cdot \delta + \ldots) = \lambda \cdot \delta + o(\delta) \quad (12.4.2.1)$$

- probabilitatea de a avea mai mult de o sosire în intervalul δ este neglijabilă:

$$P\left[N_a(\delta) = k\right] = o(\delta), \text{ cu } k \geq 2 \quad (12.4.2.2)$$

- probabilitatea ca în intervalul δ să se servească un client este semnificativă, depinzând de rata de servire, μ, şi de lungimea intervalului:

$$P\left[N_a(\delta) = 1\right] = P(\tau \leq \delta) = \mu \cdot \delta + o(\delta) \quad (12.4.2.3)$$

- probabilitatea de a avea mai mult de o plecare în intervalul δ este neglijabilă:

$$P\left[N_d(\delta) = k\right] = o(\delta), \text{ cu } k \geq 2 \quad (12.4.2.4)$$

- probabilitatea ca în intervalul δ să avem simultan două evenimente diferite, anume o sosire şi o plecare, este neglijabilă:

$$P\left[N_a(\delta) = 1, \tau \leq \delta\right] = P\left[N_a(\delta) = 1\right] \cdot P\left[\tau \leq \delta\right] = o(\delta) \quad (12.4.2.5)$$

Conform celor precizate mai sus, rezultă că procesul $N(t)$ admite doar tranziţii adiacente, fiind deci un proces de naştere şi moarte, a cărui diagramă de tranziţii este trasată în figura 12.4.3. Se scriu următoarele ecuaţii de balans:

- echilibru local (suprafaţa Σ_0), pentru $j = 0$:

$$\lambda \cdot p_0 = \mu \cdot p_1 \quad (12.4.2.6)$$

- echilibru global (suprafaţa Σ_1), pentru orice $j \geq 1$:

$$(\lambda+\mu)\cdot p_j = \lambda \cdot p_{j-1} + \mu \cdot p_{j+1} \qquad (12.4.2.7)$$

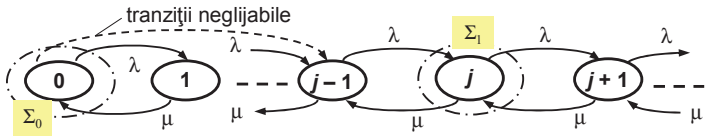


Figura 12.4.3: Diagrama ratelor de tranziţii – M/M/1

Considerarea prin recurenţă a ecuaţiilor de echilibru permite scrierea relaţiei de calcul al probabilităţii oricărei stări, pe baza cunoaşterii probabilităţii stării vide a procesului, şi anume:

$$p_j = P[N(t)=j] = \lambda/\mu \cdot p_{j-1}, \text{ adică } p_j = \rho^j \cdot p_0 \qquad (12.4.2.8)$$

în care $\rho = \lambda/\mu$. Probabilităţile de stare astfel exprimate se introduc în relaţia de normare $\sum_{j=0}^{\infty} p_j = 1$, care este o serie convergentă numai dacă $\rho < 1$, şi rezultă că:

$$p_j = (1-\rho)\cdot\rho^j \text{ pentru } \forall\, j \geq 0 \qquad (12.4.2.9)$$

Condiţia $\rho < 1$ nu este doar o restricţie matematică, impusă de convergenţa seriei, ci ea reprezintă o cerinţă reală a funcţionării unui sistem uni-server. Dacă $\rho > 1$, ar însemna că sosirea în sistem a clienţilor este mai rapidă decât posibilitatea servirii lor, ceea ce ar produce creşterea nelimitată a şirului de aşteptare, conducând la o situaţie de instabilitate a sistemului.

Relaţia (12.4.2.9) pune în evidenţă faptul că legea de distribuţie de ordinul 1 a numărului de clienţi într-un sistem M/M/1 este de natură **geometrică**.

Evaluarea performanţei sistemului se efectuează prin intermediul următorilor indicatori, enumeraţi în ordinea dictată de stabilirea relaţiilor corespunzătoare de calcul, prin valorificarea la maxim a formulei lui Little:

1) ***numărul mediu de clienţi în sistem***:

$$\mathrm{E}[N] = \sum_{j=0}^{\infty} j \cdot p_j = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda} \qquad (12.4.2.10)$$

2) ***numărul mediu al clienţilor din şir*** (lungimea medie a şirului de aşteptare):

$$\mathrm{E}[N_q] = \sum_{j=1}^{\infty} (j-1)\cdot p_j = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)} \qquad (12.4.2.11)$$

3) ***media timpului petrecut de un client în sistem*** se obţine cu formula lui Little ca fiind:

$$\mathrm{E}[T] = \frac{\mathrm{E}[N]}{\lambda} = \frac{1}{\lambda}\cdot\frac{\rho}{1-\rho} = \frac{\mathrm{E}[\tau]}{1-\rho} = \frac{1}{\mu-\lambda} \qquad (12.4.2.12)$$

4) ***media timpului de aşteptare*** se poate calcula cu ajutorul formulei lui Little, $\mathrm{E}[W] = \mathrm{E}[N_q]/\lambda$, sau ca o diferenţă între timpul de staţionare şi cel de servire, $\mathrm{E}[W] = \mathrm{E}[T] - \mathrm{E}[\tau]$, ambele variante conducând la relaţia:

$$\mathrm{E}[W] = \frac{\rho}{\mu-\lambda} \qquad (12.4.2.13)$$

5) ***traficul servit*** sau ***gradul de utilizare a unităţii de servire*** *(unice)*:

$$g = \mathrm{E}[N_s] = \lambda \cdot \mathrm{E}[\tau] = \frac{\lambda}{\mu} = \rho \qquad (12.4.2.14)$$

6) ***probabilitatea de aşteptare***:

$$P[W>0] = P[N \geq 1] = \sum_{j=1}^{\infty} p_j = 1 - p_0 = \rho \qquad (12.4.2.15)$$

7) ***productivitatea γ a sistemului*** (reprezintă frecvenţa medie cu care sistemul eliberează clienţii al căror serviciu este finalizat):

$$\gamma = \lambda_{ef} = \sum_{j=1}^{\infty} \mu_j \cdot p_j = \mu \cdot (1-p_0) = \lambda \qquad (12.4.2.16)$$

În figura 12.4.4 este prezentată curba de variaţie cu frecvenţa sosirilor a numărului mediu de clienţi prezenţi în sistem. Se poate observa că atâta timp cât frecvenţa este redusă, clienţii sunt puţin numeroşi (unul singur putând fi în serviciu), iar şirul de aşteptare destul de scurt, dar trecerea spre saturare este foarte rapidă, ea apărând odată ce frecvenţa depăşeşte o anumită limită ($\lambda_{ef} > 0{,}75\mu$).

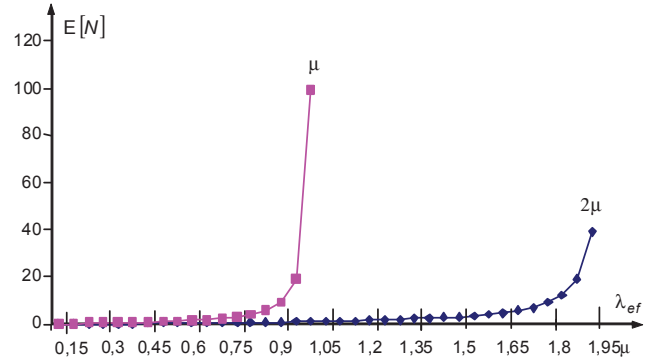


Figura 12.4.4: Variaţia cu frecvenţa sosirilor a numărului de clienţi din sistem

O metodă posibilă de extindere a domeniului de utilizare a unui sistem uniserver, în condiţii acceptabile de încărcare, este creşterea ratei de servire a clienţilor. În figura 12.4.4 este prezentată o asemenea soluţie, când prin dublarea ratei de servire se obţine o amplificare corespunzătoare a domeniului permis al ratei de sosire a clienţilor (de exemplu $\lambda_{ef} \le 1{,}8\mu$).

Aplicaţia 12.4.3

Un sistem M/M1 este alimentat de la o sursă externă cu un trafic de intensitate $a = 0{,}75$ E, clienţii sosind în medie câte unul la fiecare 4 secunde. Cum trebuie să se modifice rata de servire a serverului astfel încât timpul de aşteptare în sistem să se înjumătăţească?

Rezolvare: în situaţia iniţială durata medie de servire este $\tau = 1/\mu = 3\text{sec}$ pentru care corespunde un timp mediu de aşteptare $\mathrm{E}[W] = 9\text{sec}$. Pentru a satisface cerinţa de înjumătăţire a acestei valori trebuie aflată valoarea noii rate de servire μ' prin rezolvarea ecuaţiei: $4{,}5 = \lambda/\mu'/(\mu' - \lambda)$.

Se obţine $\mu' = 0{,}39\,\text{clienţi/sec}$, ceea ce, aşa cum era de aşteptat, corespunde unui timp de servire mai mic decât în varianta iniţială, şi anume $\tau' = 2{,}55\text{sec}$ (adică o servire mai rapidă). Trebuie, de asemenea, remarcat că prin creşterea vitezei de servire s-a modificat în consecinţă şi valoarea traficului, anume $a' = 0{,}56$ E.

12.5 Modelul M#/M/1 cu control de flux

Există sisteme de servire cu aşteptare care, în scopul evitării situaţiilor de saturare, execută un control asupra sosirii clienţilor, prin intermediul unui component suplimentar, numit *"controler"*. În acest fel, se pot înlătura situaţiile de ieşire din echilibru statistic, proprii de exemplu modelului M/M/1, la care, pentru $\rho \ge 1$, acumularea de clienţi capătă un caracter dominant. Prin acţiunea controlerului se obţine totdeauna o îmbunătăţire a performanţelor, dar cu preţul creşterii complexităţii sistemului. Asemenea sisteme, deşi mai costisitoare, sunt în general necesare în cadrul reţelelor de comunicaţii pentru a realiza regularizarea fluxurilor informaţionale şi asigurarea unor performanţe constante de-a lungul reţelelor. În consecinţă, indiferent de "politica" urmată în funcţionarea sa, controlerul trebuie să aibă ca obiectiv obţinerea unor rate de acces în sistem care să fie dependente de starea sistemului.

Pentru a demonstra eficienţa sistemelor cu accesul controlat al clienţilor se propune analiza unui ***model cu descurajarea clienţilor***, în care rata de acces în sistem scade proporţional cu numărul clienţilor deja prezenţi, adică:

$$\lambda_j = \lambda/(j+1) \qquad (12.5.1)$$

în care λ este rata medie cu care o sursă infinit de mare oferă clienţii într-un proces de tip Poisson. În plus se acceptă că timpii de serviciu au o distribuţie

exponenţial negativă de medie $1/\mu$ şi că sistemul este echipat cu un şir infinit şi cu un singur server.

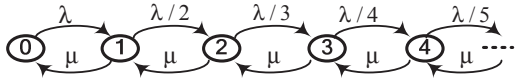


Figura 12.5.1: Evoluţia procesului M#/M/1

Înseamnă că, în conformitate cu clasificarea Kendall, un asemenea sistem poate fi numit M#/M/1, semnul # marcând faptul că rata procesului Poisson, de sosire, este controlată de starea sistemului.

Evoluţia procesului de prelucrare ce corespunde unui asemenea sistem este reprezentată în figura 12.5.1. Valoarea probabilităţii unei stări particulare a sistemului se calculează cu formula:

$$p_j = p_0 \prod_{k=0}^{j-1} \frac{\lambda_k}{\mu_{k+1}} = p_0 \frac{\rho^j}{j!} \qquad (12.5.2)$$

ceea ce face ca, prin intermediul condiţiei de normare, să se obţină:

$$p_j = \frac{\rho^j}{j!} \cdot \mathrm{e}^{-\rho} \qquad (12.5.3)$$

O primă observaţie referitoare la sistemul M#/M/1 este aceea că *nu mai există nici o restricţie cu privire la valoarea traficului oferit sistemului*, teoretic el putând fi oricât de mare. Sosirea însă în sistem a clienţilor este regularizată prin mărimea ratei intrărilor, care este dependentă de starea sistemului. Se poate calcula valoarea medie a ratei de intrare, λ_{ef}, în condiţiile serviciului de rată μ independentă de stare, şi se obţine că:

$$\lambda_{ef} = \sum_{j=0}^{\infty} \lambda_j \cdot p_j = (1 - p_0) \cdot \mu = (1 - \mathrm{e}^{-\rho}) \cdot \mu \qquad (12.5.4)$$

Evaluarea performanţei se efectuează cu ajutorul următorilor indicatori:

1) ***numărul mediu de clienţi prezenţi în sistem*** **:**

$$\mathrm{E}[N] = \rho \qquad (12.5.5)$$

Dacă se compară acest rezultat cu valoarea corespunzătoare unui sistem clasic M/M/1, valoare dată de relaţia (12.5.5), atunci se poate confirma şi matematic faptul că în sistemul cu acces controlat sunt mai puţini clienţi prezenţi, deci cu siguranţă este necesară o memorie de capacitate mai mică (din cei prezenţi un client este servit, restul fiind în aşteptare).

2) ***timpul mediu de staţionare în sistem*** **:**

$$\mathrm{E}[T] = \frac{\mathrm{E}[N]}{\lambda_{ef}} = \frac{\rho}{\mu \cdot (1 - \mathrm{e}^{-\rho})} \qquad (12.5.6)$$

Se pot trage concluziile următoare:

- pentru valori foarte mici ale sarcinii oferite ($\rho \to 0$) timpul mediu de

staţionare este practic egal cu timpul mediu de servire, ca şi când nu ar exista aşteptare;

- pentru valori mari ale sarcinii oferite ($\rho \to \infty$) timpul mediu de staţionare creşte practic liniar cu sarcina.

3) ***lungimea medie a şirului de aşteptare :***

$$\mathrm{E}\left[N_q\right] = \sum_{j=1}^{\infty} (j-1)\cdot p_j = \rho - 1 + \exp(-\rho) \qquad (12.5.7)$$

Aplicaţia 12.8

Determinaţi relaţia de calcul al traficului servit corespunzătoare modelului M#/M/1. Comparaţi sistemele M/M/1 şi M#/M/1 din punctul de vedere al traficului servit.

12.6 Modelul M/M/1/∞/r/FIFO

Acest model poate fi considerat ca un caz particular al celui anterior, dar el este în sine un model frecvent utilizat în aplicaţiile practice, liniile de transmisiuni între diferite noduri de comunicaţie fiind în realitate sisteme uniserver cu cozi finite.

Aplicaţia 12.6.2

Determinaţi probabilităţile de stare ce caracterizează modelul M/M/1/∞/r/FIFO.

Indicaţie: Se poate folosi graful de tranziţii din figura 12.4.3, dar cu observaţia esenţială că el se aplică pentru un număr finit de stări, determinat de dimensiunea r a sistemului. În consecinţă relaţia de normare reprezintă în acest caz o serie finită, ceea ce nu impune restricţii privind raportul dintre mărimile λ şi μ. În plus, cazul $\rho = 1$ impune o tratare separată, fiind necesară rescrierea sumei probabilităţilor de stare în mod corespunzător. Se obţine astfel:

- pentru $\rho \neq 1$:
$$\begin{cases} p_0 = \dfrac{1-\rho}{1-\rho^{r+1}} \\ p_j = \dfrac{1-\rho}{1-\rho^{r+1}} \cdot \rho^j \end{cases} \qquad (12.6.12)$$

- pentru $\rho = 1$: $p_0 = p_j = (r+1)^{-1}$ (12.6.13)

Indicatorii pentru evaluarea performanţei sunt:

1) ***numărul mediu de clienţi prezenţi în sistem:***

$$\mathrm{E}[N] = \sum_{j=0}^{r} j\cdot p_j = \begin{cases} \dfrac{\rho}{1-\rho} - \dfrac{(r+1)\,\rho^{r+1}}{1-\rho^{r+1}} & \text{pentru } \rho \neq 1 \\ r/2 & \text{pentru } \rho = 1 \end{cases} \qquad (12.6.14)$$

2) ***numărul mediu de clienţi în aşteptare*** (lungimea medie a şirului):

$$\mathrm{E}\left[N_q\right] = \sum_{j=1}^{r} (j-1)\cdot p_j = p_0 \cdot \rho \sum_{j=1}^{q} \rho^{j} = p_0 \cdot \rho^2 \frac{1 - r\rho^{r-1} + (r-1)\rho^r}{(1-\rho)^2} \qquad (12.6.15)$$

3) ***probabilitatea de aşteptare:***

$$P[W > 0] = \sum_{j=1}^{r-1} p_j = 1 - p_0 - p_r = \rho \frac{1-\rho^{r-1}}{1-\rho^{r+1}} \qquad (12.6.16)$$

4) ***probabilitatea de blocare:***

$$p_b = p_r = p_0 \cdot \rho^r = \begin{cases} \rho^r \dfrac{1-\rho}{1-\rho^{r+1}} & \text{pentru } \rho \neq 1 \\ 1/(r+1) & \text{pentru } \rho = 1 \end{cases} \qquad (12.6.17)$$

Variaţia probabilităţii de blocare cu traficul oferit, ρ, este prezentată în figura 12.6.3, **domeniul de blocare** (de congestie a sistemului) fiind intervalul în care traficul oferit impune sistemului o funcţionare în vecinătatea umplerii complete, însoţită de o respingere semnificativă de clienţi. Din curbele prezentate se poate observa că probabilitatea blocării este continuu crescătoare cu traficul oferit, fiind cu atât mai pronunţată cu cât şirul este de mărime mai mică.

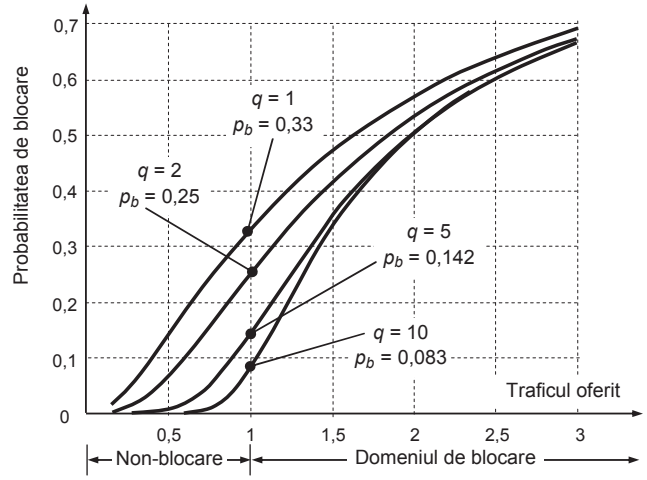


Figura 12.6.3: Variaţia probabilităţii de blocare cu traficul oferit la diverse şiruri

5) ***productivitatea sistemului:***

Aşa cum se precizează şi în figura 12.6.1, sistemul în discuţie respinge cu o rată medie λ_L clienţii care sosesc în momentele sale de blocare. Întrucât, blocarea reprezintă de fapt starea de încărcare completă a sistemului, înseamnă că *rata medie* λ_{ef} *de tranzit prin sistem* este:

$$\lambda_{ef} = \lambda - \lambda_L = \lambda \cdot (1 - p_b) \qquad (12.6.18)$$

Cum însă fiecare client acceptat în sistem urmează şi să fie servit (imediat sau după o aşteptare oarecare), înseamnă că productivitatea este dată de relaţia:

$$\gamma = \lambda_{ef} = \lambda \cdot \frac{1-\rho^r}{1-\rho^{r+1}} \qquad (12.6.19)$$

Expresia ratei medii de părăsire a sistemului (a productivităţii) poate fi obţinută şi prin medierea ratelor de servire din toate stările posibile, adică:

$$\gamma = \lambda_{ef} = 0 \cdot P[N(t)=0] + \mu \cdot P[N(t)>0]$$

cu: $P[N(t)=0] = p_0$ (probabilitatea ca în sistem să nu fie nici un client),

$P[N(t)>0] = 1 - p_0$ (probabilitatea ca în sistem să fie cel puţin un client).

Deci: $\lambda_{ef} = 0 + \mu(1-p_0) = \mu(1-p_0)$ şi făcând calculele, pentru $\rho \neq 1$, rezultă expresia:

$$\frac{\lambda_{ef}}{\mu} = \frac{\rho \cdot (1-\rho^r)}{1-\rho^{r+1}} \qquad (12.6.20)$$

ce este similară cu expresia anterioară (12.6.19). Din această exprimare, se poate însă deduce mai simplu că rata maximă de tranzit prin sistem tinde către rata μ de servire a serverului, când $\lambda \to \infty$ (traficul oferit devine foarte mare).

Ţinând seama că rata de tranzit prin sistem este, în acest caz, diferită de aceea a cererilor de serviciu, λ, trebuie făcută o diferenţiere clară între următoarele elemente:

- ***traficul oferit*** sau ***sarcina oferită***, ce se calculează prin produsul:

$$\lambda_{[\text{clienţi/sec}]} \cdot \mathrm{E}[\tau]_{[\text{secunde de servire/client}]}$$

- ***traficul servit*** sau ***sarcina servită***, ce se evaluează prin produsul:

$$\lambda_{ef\,[\text{clienţi/sec}]} \cdot \mathrm{E}[\tau]_{[\text{secunde de servire/client}]}$$

Rata de tranzit prin sistem corespunde *clienţilor acceptaţi* de sistem, ceea ce permite utilizarea relaţiei lui Little pentru stabilirea următorilor indicatori de performanţă:

6) ***timpul mediu petrecut de un client în sistem***:

$$\mathrm{E}[T] = \frac{\mathrm{E}[N]}{\lambda_{ef}} \qquad (12.6.21)$$

7) ***timpul mediu de aşteptare***:

$$\mathrm{E}[W] = \frac{\mathrm{E}[N_q]}{\lambda_{ef}} \qquad (12.6.22)$$

8) ***gradul de utilizare a unităţii de servire***:

$$g = \lambda_{ef} \cdot \mathrm{E}[\tau] \qquad (12.6.23)$$

Pe lângă indicatorii de performanţă menţionaţi anterior, modelul în discuţie mai este caracterizat şi de:

9) ***probabilitatea respingerii unui client***, p_L. Pentru stabilirea formulei sale de calcul se pleacă de la faptul că respingerea unui client este echivalentă cu faptul că la sosirea acestuia sistemul este blocat. Prin urmare, este adevărată următoarea egalitate:

$$p_L = P[N(t) = r \,|\, N_a(\delta) = 1] \qquad (12.6.24)$$

unde δ este un interval de timp suficient de mic (teoretic $\delta \to 0$), astfel încât pe parcursul său, dacă apare o sosire, $N_a(\delta) = 1$, să nu se mai întâmple şi alte evenimente, de genul sosiri sau serviri (condiţie care este satisfăcută ţinând cont că procesul analizat aparţine clasei proceselor de naştere şi moarte; vezi discuţia de la modelul M/M/1).

Fiind o probabilitate condiţionată, pentru probabilitatea de respingere se pot scrie următoarele relaţii:

$$P[N(t)=k \,|\, N_a(\delta)=1] = \frac{P[N(t)=r, N_a(\delta)=1]}{P[N_a(\delta)=1]} = \frac{P[N_a(\delta)=1 \,|\, N(t)=r] \cdot P[N(t)=r]}{P[N_a(\delta)=1]}$$

în care, dat fiind faptul că starea sistemului nu influenţează procesul de sosire a clienţilor, probabilitatea sosirii unui client atunci când sistemul este blocat este identică cu probabilitatea sosirii unui client indiferent de starea acestuia, în consecinţă, atât pentru modelul M/M/1/∞/r, cât şi pentru oricare alt model care prezintă această proprietate, probabilitatea de pierdere este egală ca valoare cu probabilitatea blocării sistemului:

$$p_L \overset{\text{M/M/1/}\infty\text{/r}}{=} P[N(t)=r] = p_b \qquad (12.6.25)$$

12.7 Modelul M/M/s

În cazul modelului multi-server **M/M/s** intrarea clienţilor este de tip Poisson, iar servirile sunt exponenţiale, cu aceeaşi rată pentru toate cele s servere din dotare. Analiza acestui modelul poate pleca prin a considera relaţia (10.4.6), valabilă pentru orice proces de naştere şi moarte, şi care permite determinarea probabilităţii oricărei stări pe baza probabilităţii stării vide şi a ratelor de tranziţie între stări corespunzătoare cazului de faţă.

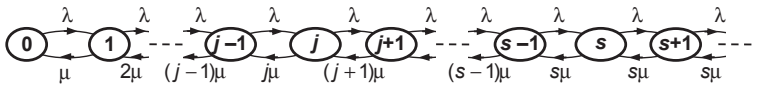


Figura 12.7.1: Diagrama ratelor de tranziţii – M/M/s

Aşa cum se precizează şi în diagrama din figura 12.7.1, în ceea ce priveşte rata de sosire a clienţilor, ea este independentă de stare, adică:

$$\lambda_j = \lambda \text{, pentru orice } j \geq 0 \qquad (12.7.1)$$

Dacă în sistem sunt prezenţi mai mulţi clienţi decât numărul s al serverelor, ceea ce înseamnă că sistemul este într-o stare $j \geq s$, atunci evident că toate serverele sunt ocupate, fiecare având o rată medie de servire μ, iar sistemul prezintă o rată medie $s \cdot \mu$ de eliberare a clienţilor. Dacă însă în sistem sunt mai puţini clienţi decât numărul serverelor, adică dacă sistemul este într-o stare $j < s$, atunci doar *j* servere din *s* sunt ocupate şi eliberarea clienţilor se face cu o rată medie $j \cdot \mu$. Deci, concentrând cele spuse:

$$\mu_j = \begin{cases} j \cdot \mu & \text{pentru } 1 \leq j < s \\ s \cdot \mu & \text{pentru } j \geq s \end{cases} \qquad (12.7.2)$$

Argumentarea matematică a expresiilor conţinute în relaţia (12.7.2) se poate dezvolta în modul următor: se presupune că la un moment dat sunt ocupate j resurse, timpul, X, până când se va încheia servirea unuia din clienţii deserviţi de ele, în condiţiile menţinerii stării până în acel moment, este dat de relaţia:

$$X = \min\left(\tau_1, \tau_2, \cdots, \tau_j\right)$$

unde, τ_i sunt variabile aleatorii i.i.d după o lege exponenţială de parametru μ. În acest caz, determinarea funcţiei de distribuţie a variabilei aleatorii X presupune raţionamentele sugerate de următoarele relaţii:

$$P[X \leq t] = 1 - P[X > t];$$

$$P[X > t] = P\left[\min\left(\tau_1, \tau_2, \cdots, \tau_j\right) > t\right] = P\left[\tau_1 > t, \tau_2 > t, \cdots, \tau_j > t\right] \overset{\text{independente}}{=}$$

$$= P[\tau_1 > t] \cdot P[\tau_2 > t] \cdot \ldots \cdot P\left[\tau_j > t\right] \overset{\text{identic distribuite}}{=} \left\{ P[\tau > t] \right\}^j = \left(\mathrm{e}^{-\mu \cdot t} \right)^j$$

şi deci: $P[X \leq t] = 1 - \exp(-j \cdot \mu t)$. Concluzia finală este că variabila analizată, X, urmează tot o distribuţie exponenţială de medie $j \cdot \mu$.

Făcând apel la (10.4.6) şi folosind exprimările precizate anterior pentru ratele de transfer între stări, în sens ascendent şi respectiv descendent al tranziţiilor, se obţine că:

$$p_j = \begin{cases} p_0 \cdot a^j / j! & \text{pentru } j < s \\ p_0 \cdot a^s / s! \cdot \rho^{j-s} & \text{pentru } j \geq s \end{cases} \qquad (12.7.3)$$

în care s-a notat **utilizarea serverului** cu $\rho = a/s = \lambda/(s\mu)$.

Relaţia de normare, aplicată unui număr infinit de stări, permite determinarea probabilităţii stării vide şi conduce la expresia:

$$p_0 = \left(\sum_{j=0}^{s-1} a^j / j! + a^s / s! \cdot \sum_{j=s}^{\infty} \rho^{j-s} \right)^{-1} \qquad (12.7.4)$$

Se precizează că pentru a asigura situaţia de stabilitate în funcţionarea a sistemului trebuie îndeplinită condiţia $\rho < 1$, echivalentă cu $\lambda < s\mu$. Aceasta înseamnă că, din punct de vedere statistic, rata medie a sosirilor trebuie să fie mai mică decât media ratei potenţiale maxime de servire a sistemului. În aceste condiţii, seria infinită din expresia (12.7.4) este convergentă şi deci:

$$p_0 = \left(\sum_{j=0}^{s-1} \frac{a^j}{j!} + \frac{a^s}{s!} \cdot \frac{1}{1-\rho} \right)^{-1} \qquad (12.7.5)$$

Pe baza ei, probabilităţile de stare ale modelului M/M/s se calculează cu relaţia:

$$p_j = \begin{cases} p_0 \cdot a^j / j! & \text{pentru } 1 \leq j < s \\ p_0 \cdot \rho^{j-s} \cdot a^s / s! & \text{pentru } j \geq s \end{cases} \qquad (12.7.6)$$

Indicatorii de performanţă sunt, în acest caz, următorii:

1) ***probabilitatea de aşteptare***, care reprezintă probabilitatea ca un client odată sosit în sistem să aştepte în şir, fiind, deci, echivalentă cu probabilitatea ca toate resursele să fie ocupate:

$$P[W > 0] = P[N(t) \geq s] = \sum_{j=s}^{\infty} p_j = \frac{p_s}{1-\rho} \qquad (12.7.7)$$

Această relaţie reprezintă ***formula Erlang-C*** şi este notată în literatura de specialitate prin simbolurile $C(a,s)$ sau $E_2(a,s)$.

2) ***numărul mediu de clienţi din şirul de aşteptare***:

$$\mathrm{E}[N_q] = \sum_{j=s}^{\infty} (j-s) \cdot p_j = \rho/(1-\rho) \cdot C(a,s) \qquad (12.7.8)$$

3) ***timpul mediu de aşteptare***:

$$\mathrm{E}[W] = \frac{\mathrm{E}[N_q]}{\lambda} = \frac{1}{\mu s} \cdot \frac{C(a,s)}{1-\rho} \qquad (12.7.9)$$

4) ***timpul mediu de staţionare în sistem:***

$$\mathrm{E}[T] = \mathrm{E}[W] + \mathrm{E}[\tau] = \frac{1}{\mu} \cdot \left(1 + \frac{C(a,s)}{s-a} \right) \qquad (12.7.10)$$

5) ***numărul mediu de servere ocupate***, care este egal cu numărul mediu al clienţilor serviţi şi care constituie media ***traficului servit***:

$$\mathrm{E}[N_s] = \sum_{j=0}^{s} j \cdot p_j + \sum_{j=s+1}^{\infty} s \cdot p_j = a \qquad (12.7.11)$$

Rezultatul este evident, deoarece nu există în acest caz vreun "eşec" de acceptare în sistem, toţi clienţii oferiţi sistemului fiind serviţi, imediat sau după o oarecare aşteptare.

6) ***gradul de utilizare a serverelor*** :

$$g = \mathrm{E}[N_s]/s = a/s \qquad (12.7.12)$$

7) ***numărul mediu de clienţi prezenţi în sistem***:

$$\mathrm{E}[N] = \lambda \cdot \mathrm{E}[T] = \mathrm{E}[N_q] + \mathrm{E}[N_s] \qquad (12.7.13)$$

Aplicaţie 12.7.1

Luând în considerare indicatorii de performanţă care le caracterizează, realizaţi o analiză comparativă a modelelor M/M/2 şi M/M/1 în cazul în care ratele de servire ale serverelor sunt μ, respectiv 2μ.

Indicaţie: Comparaţia se face pe baza curbelor de variaţie a indicatorilor de performanţă urmărită, ce se trasează pe acelaşi sistem de axe pentru fiecare indicator în parte (vezi figura 12.7.2). Astfel, analizând, de exemplu, timpul mediu de staţionare petrecut de un client în sistem, $\mathrm{E}[T]$, normat cu durata medie de servire, $1/\tau$, se pune în evidenţă faptul că modelul M/M/1 înzestrat cu un server de două ori mai rapid decât serverele modelului M/M/2 este mai eficient decât acesta din urmă care, însă, îi este superior prin redundanţa şi costurile reduse de exploatare care îl caracterizează.

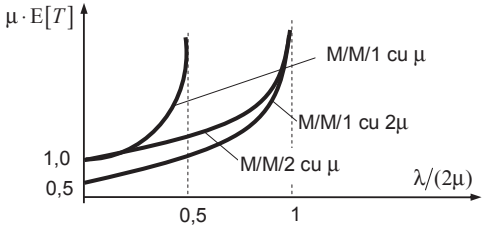


Figura 12.7.2: Variaţia cu traficul oferit a timpului normat de staţionare

12.8 Modelul M/M/s/∞/s cu pierderi

Acest model este adecvat analizei sistemelor de telecomunicaţii care nu dispun de mijloace şi metode de menţinere a clienţilor în aşteptare în momentele în care toate serverele sunt ocupate. Înseamnă că numărul de poziţii permise în sistem coincide cu dimensiunea s a grupei de servere, iar timpul de tranzitare a

sistemului este chiar timpul de servire a oricărui client acceptat în sistem. Este cazul aşa numitelor ***sisteme cu pierderi***, ce resping toţi clienţii care încearcă pătrunderea în sistem când acesta este blocat (toate resursele sunt ocupate). Diagrama ratelor de tranziţie pentru modelul M/M/s/∞/s este precizată în figura 12.8.1 şi, evident, se referă la un număr finit de stări, şi anume $s+1$, ratele de tranziţie fiind:

$$\text{pentru } 0 \le j \le s \quad \begin{cases} \lambda_j = \lambda \\ \mu_j = j \cdot \mu \end{cases} \qquad (12.8.1)$$

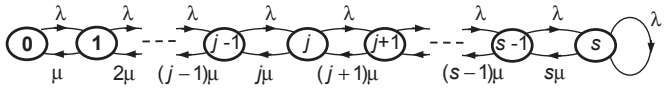


Figura 12.8.1: Diagrama ratelor de tranziţii pentru modelul M/M/s/∞/s

În figura 12.8.2 este ilustrată structură de principiu a unui sistem ce funcţionează conform acestui model. Sunt evidenţiate tipurile de volume de trafic ce intervin: traficul oferit de natură întâmplătoare, $X(t)$, şi volumele rezultate după prelucrare, scurs $Y(t)$ şi respins $R(t)$, ambele fiind de natură deterministă.

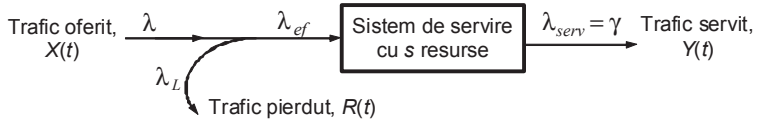


Figura 12.8.2: Sistem de prelucrare cu pierderi

Plecând de la expresia general valabilă (10.4.6), aplicând-o pentru cele $s+1$ stări şi folosind relaţia de normare se obţine că:

$$p_0 = \left(\sum_{j=0}^{s} \frac{a^j}{j!} \right)^{-1} \text{ şi } p_j = p_0 \cdot \frac{a^j}{j!}, \text{ cu } a = \frac{\lambda}{\mu}. \qquad (12.8.2)$$

Indicatorii de performanţă ai modelului M/M/s/∞/s sunt:

1) ***probabilitatea de blocare*** este probabilitatea ca toate resursele (adică serverele) să fie ocupate şi este dată de ***formula Erlang-B***, care se notează prin $B(a,s)$ sau $E_1(a,s)$:

$$p_b = P[N(t) = s] = p_s = \frac{a^s}{s!} \Big/ \sum_{j=0}^{s} \frac{a^j}{j!} \qquad (12.8.3)$$

2) ***productivitatea sistemului***, γ, este identică cu rata efectivă de intrare a clienţilor în sistem, λ_{ef}:

$$\gamma = \lambda_{ef} = \lambda \cdot [1 - B(a,s)] \qquad (12.8.4)$$

Diferenţa până la rata λ, cu care sursa oferă clienţi, reprezintă ***rata de respingere a clienţilor*** sau rata de pierdere, $\lambda_L = \lambda \cdot B(a,s)$.

3) ***numărul mediu de clienţi în sistem*** (intensitatea medie de trafic servit):

$$\mathrm{E}[N] = \sum_{j=0}^{s} j \cdot p_j = \lambda_{ef} \cdot \mathrm{E}[\tau] = a \cdot [1 - B(a,s)] = \mathrm{E}[Y(t)] \quad (12.8.5)$$

4) ***gradul de utilizare a serverelor*** :

$$g = \lambda_{ef} \cdot \mathrm{E}[\tau] / s = \mathrm{E}[Y(t)] / s \quad (12.8.6)$$

Aplicaţia 12.8.1

Plecând de la formula Erlang-B, verificaţi corectitudinea relaţiei de recurenţă ce urmează:

$$B(a,s) = \frac{a \cdot B(a,s-1)}{s + a \cdot B(a,s-1)} \quad (12.8.7)$$

* * *

Aplicaţia 12.8.2

Aplicând relaţia de recurenţă (12.8.7), stabiliţi numărul unităţilor de servire ale unui sistem M/M/s/∞/s astfel încât probabilitatea de blocare să fie de cel mult 10^{-3}, pentru un trafic oferit de intensitate egală cu $2\,[\mathrm{E}]$.

12.10 Modelul M/M/s/∞/r/FIFO

Modelul M/M/s/∞/r/FIFO este cu siguranţă mai aproape de situaţiile reale, decât modelul M/M/s, pentru că, aşa cum este el definit conform clasificării Kendall, ia în considerare o capacitate finită de stocare, $q = r - s$. Deci cel mult un număr q de clienţi pot fi acceptaţi în sistem în poziţiile de aşteptare, urmând ca ei să fie serviţi, după o disciplină FIFO, şi să fie eliberaţi apoi după servire cu o rată λ_{ef}. Toţi ceilalţi clienţi, care sunt oferiţi de către sursa infinită, cu rata medie λ independentă de starea sistemului, dar care sosesc după completarea capacităţii de stocare vor fi respinşi de sistem, cu o rată λ_L (figura 12.10.1), constituind pierderi. Această ultimă situaţie corespunde stării de *blocare* (umplere completă) a sistemului, ce înglobează deopotrivă ocuparea serverului precum şi a şirului de aşteptare.

Rezultă, deci, că modelul M/M/s/∞/r/FIFO caracterizează aşa numitele ***sisteme cu aşteptare şi pierderi***, pentru care va trebui evaluat şi un indicator

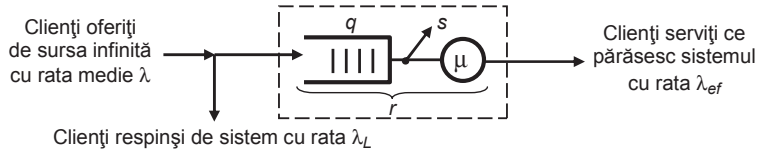


Figura 12.10.1: Structura generală a unui sistem M/M/s/∞/r/FIFO

suplimentar de performanţă, anume ***probabilitatea de pierderi***, care caracterizează comportamentul sistemului în starea sa finală, când toate cele $r = q + s$ poziţii prevăzute în sistem sunt ocupate şi nici un alt client nou nu mai poate fi acceptat.

Din punct de vedere matematic, diferenţa faţă de modelul anterior M/M/s apare datorită faptului că aici există doar un număr finit de stări, anume $s + q + 1$, iar starea finală $r = s + q$ caracterizând fenomenul de respingere a clienţilor (apariţia pierderilor).

Parametrii procesului de prelucrare sunt:

$$\lambda_j = \begin{cases} \lambda & \text{pentru } 0 \le j \le r-1 \\ 0 & \text{pentru } j = r \end{cases}$$
$$\mu_j = \begin{cases} j \cdot \mu & \text{pentru } 0 \le j < s \\ s \cdot \mu & \text{pentru } s \le j \le r \end{cases} \quad (12.10.1)$$

Expresiile probabilităţilor de stare sunt aceleaşi ca şi în cazul modelului M/M/s, adică:

$$p_j = \begin{cases} p_0 \cdot a^j / j! & \text{pentru } 1 \le j < s \\ p_0 \cdot \rho^{j-s} \cdot a^s / s! & \text{pentru } s \le j \le r \end{cases} \quad (12.10.2)$$

dar **probabilitatea stării vide** se obţine prin aplicarea relaţiei de normare pe numărul finit de stări, deci:

$$p_0 = \left(\sum_{j=0}^{s-1} \frac{a^j}{j!} + \frac{a^s}{s!} \sum_{j=s}^{r} \rho^{j-s} \right)^{-1} = \left(\sum_{j=0}^{s-1} \frac{a^j}{j!} + \frac{a^s}{s!} \cdot \frac{1-\rho^{q+1}}{1-\rho} \right)^{-1} \quad (12.10.3)$$

Indicatorii de performanţă sunt, în acest caz, următorii:

1) ***probabilitatea de aşteptare*** – reprezintă probabilitatea ca un client nou oferit când toate serverele sunt ocupate să poată fi acceptat în şirul de aşteptare:

$$p_w = P[W > 0] = P[s \le N(t) \le r-1] = \sum_{j=s}^{r-1} p_j = p_0 \cdot \frac{a^s}{s!} \cdot \frac{1-\rho^q}{1-\rho} \quad (12.10.4)$$

2) ***probabilitatea de pierderi*** – reprezintă probabilitatea ca un client nou să fie respins de sistem, deoarece nu mai este disponibilă vreo poziţie nici la servire nici la aşteptare; este de fapt probabilitatea stării finale:

$$p_L = p_b = p_0 \cdot \frac{a^s}{s!} \cdot \rho^q \quad (12.10.5)$$

pe baza căreia poate fi evaluată rata de acceptare a clienţilor în sistem:

$$\lambda_{ef} = \lambda(1 - p_L) \quad (12.10.6)$$

3) ***numărul mediu de clienţi din şirul de aşteptare***:

$$\mathrm{E}\left[N_q\right] = \sum_{j=s}^{r} (j-s)\cdot p_j = p_0\, a^s / s! \cdot \sum_{k=1}^{q} k \cdot \rho^k \qquad (12.10.7)$$

4) ***numărul mediu de servere ocupate*** – este egal cu numărul mediu al clienţilor serviţi şi constituie media ***traficului servit***:

$$\mathrm{E}\left[N_s\right] = \sum_{j=0}^{s} j \cdot p_j + \sum_{j=s+1}^{r} s \cdot p_j = a \cdot (1 - p_L) \qquad (12.10.8)$$

În partea finală a expresiei s-a precizat faptul că volumul traficului servit este mai mic decât cel al traficului oferit, diferenţa fiind evident datorată traficului pierdut (respins) din lipsă de resurse.

5) ***timpul mediu de aşteptare***:

$$\mathrm{E}[W] = \mathrm{E}\left[N_q\right] / \lambda_{ef} \qquad (12.10.9)$$

6) ***timpul mediu de staţionare în sistem:***

$$\mathrm{E}[T] = \mathrm{E}[W] + \mathrm{E}[\tau] \qquad (12.10.10)$$

7) ***numărul mediu de clienţi prezenţi în sistem***:

$$\mathrm{E}[N] = \mathrm{E}\left[N_q\right] + \mathrm{E}\left[N_s\right] = \lambda_{ef} \cdot \mathrm{E}[T] \qquad (12.10.11)$$

Aplicaţia 12.10.1

Analizaţi care este influenţa numărului poziţiilor de aşteptare asupra indicatorilor de performanţă.

Indicaţii: se va analiza variaţia probabilităţilor de aşteptare şi de pierderi în raport cu modificarea numărului permis al poziţiilor de aşteptare. Se vor obţine curbe de genul celor din figura 12.10.2, ce au fost trasate pentru cazul particular: $a = 3\,\mathrm{E}$ şi $s = 5$ servere.

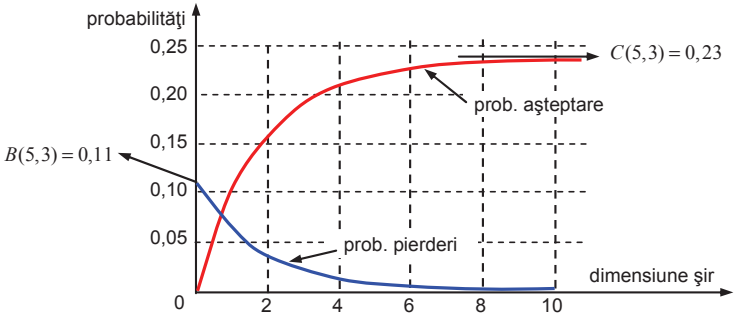


Figura 12.10.2: Influenţa numărului de poziţii de aşteptare asupra indicatorilor de performanţă

Se observă că, aşa cum era de aşteptat, pe măsură ce q creşte probabilitatea de pierderi pornind de la valoarea $B(a,s) = 0{,}1101$, ce corespunde

sistemului cu pierderi nete ($q = 0$), scade asimptotic spre zero. În schimb, probabilitatea de aşteptare, pornind de la zero, creşte asimptotic spre valoarea $C(a,s) = 0{,}2362$, ce corespunde şirului infinit.

Rezultă deci că un "sistem real", cu aşteptare şi pierderi, se situează între cele două sisteme studiate anterior: sistemul M/M/s/∞/s cu pierderi nete şi sistemul M/M/s cu aşteptare infinită. Cum indicatorii observaţi ating destul de repede valorile ce caracterizează cazul şirului infinit, înseamnă că, din punct de vedere practic, un număr limitat de poziţii de aşteptare poate fi suficient pentru a se obţine o calitate satisfăcătoare.

Pentru cazul particular ales, dacă s-ar impune ca normă de pierderi $p_L \le 10^{-3}$, atunci dimensiunea optimă a şirului ar fi $q = 9$, pentru care ar corespunde $p_a = 0{,}2341$.

12.11 Modelul M/M/1/c

Procesele de naştere şi moarte modelează, la modul general, sisteme a căror stare poate controla atât frecvenţa sosirilor, cât şi pe cea a servirilor, relaţia (10.4.6) fiind valabilă pentru astfel de situaţii, dacă se respectă anumite condiţii necesare stabilirii unui regim de funcţionare staţionar din punct de vedere probabilistic. Un caz particular, des întâlnit în realitate, îl reprezintă sistemele care deservesc o sursă cu număr finit de clienţi. Această situaţie corespunde configuraţiei din figura 12.11.1, în care un ”server” deserveşte mai mulţi "clienţi".

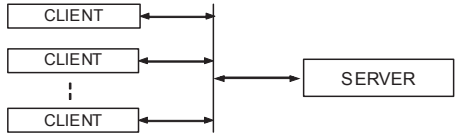


Figura 12.11.1: Exemplu de configuraţie cu mai mulţi utilizatori

În acest context, clienţii se pot găsi, din punct de vedere al sistemului, în două stări distincte:

- **IDLE** – în care clientul îşi pregăteşte o cerere de servire (tranzacţie), care presupune, de exemplu, desfăşurarea unui protocol între client şi server în vederea îndeplinirii unui serviciu,
- **BUSY** – în care clientul aşteaptă executarea serviciului cerut, cererea găsindu-se în sistem.

Modelul de analiză, asociat acestei situaţii, în cazul în care sistemul dispune de o singură unitate de servire şi capacitate de stocare suficientă pentru a înregistra numărul maxim de cereri de servire, are reprezentarea simbolică din figura 12.11.2. În privinţa proceselor de sosire şi servire a cererilor, modelul consideră că:

- durata dintre momentul încheierii unei tranzacţii şi momentul iniţierii

cererii următoare de către acelaşi terminal respectă o distribuţie exponenţială de medie $1/\alpha$,

- timpul de servire a cererilor urmează o distribuţie exponenţială de medie $1/\mu$.

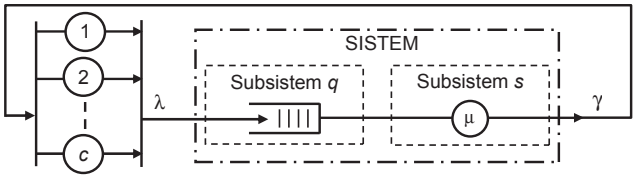


Figura 12.11.2: Structura modelului M/M/1/c

Aceste două ipoteze conduc la concluzia că procesul $N(t)$ este un lanţ Markov continuu în timp care, în plus, aparţine clasei de procese de naştere şi moarte.

Aplicaţia 12.11.1

Determinaţi expresia de calcul a ratei cererilor de servire.

Rezolvare: Generarea unei cereri de serviciu în momentul θ, inclus într-un interval infinitezimal, $(t, t+\delta)$, unde t este un moment arbitrar, de referinţă, de către un terminal aflat în starea IDLE la momentul t, are probabilitatea:

$$p = P[\theta \le \delta] = 1 - e^{-\alpha\cdot\delta} \cong \alpha\cdot\delta + o(\delta) \qquad (12.11.1)$$

Prin urmare, ansamblu surselor aflate la un moment dat t în starea IDLE respectă, din punct de vedere al generării de cereri în intervalul $(t, t+\delta)$, o distribuţie binomială de parametri: numărul terminalelor în repaus, exprimat prin diferenţa $c-k$, unde c este numărul total al terminalelor şi k este numărul terminalelor active (cu cereri în sistem), şi p, probabilitatea de generare cerere, cu formula de calcul stabilită anterior, din relaţia (12.11.1). În aceste condiţii, pe baza probabilităţilor elementare corespunzătoare, p_j, putem stabili următoarea expresie a numărului mediu de cereri generate în intervalul δ:

$$E[N_a(\delta)] = \sum_{j=0}^{c-k} j\cdot p_j = (c-k)\cdot\alpha\cdot\delta$$

Aproximând distribuţia binomială cu o distribuţie Poisson, înseamnă că rata cererilor de servire a unui număr de $c-k$ surse aflate în starea IDLE este dată de relaţia:

$$\lambda_k = (c-k)\cdot\alpha \qquad (12.11.2)$$

relaţie care pune în evidenţă dependenţa acesteia de starea sistemului.

* * *

Aplicaţia 12.11.2

Verificaţi apartenenţa procesului considerat la clasa proceselor de naştere şi moarte.

Indicaţie: Se determină dacă, indiferent de starea procesului, $N(t)=k$, numai tranziţiile adiacente au probabilităţi semnificative. Acest demers presupune analiza a cinci cazuri care sunt prezentate succint în continuare:

- **Probabilitatea ca** pe un interval infinitezimal, δ, **doar o sursă** inactivă **să genereze o cerere.** Formula de calcul a acestei probabilităţi este:

$$P[N_a(\delta)=1 \mid N(t)=k] = C_{c-k}^{1}\cdot P[\theta\le\delta]\cdot P[\theta>\delta]^{c-k-1} \cong (c-k)\cdot\alpha\cdot\delta + o(\delta) \qquad (12.11.3)$$

relaţie care evidenţiază o valoare semnificativă a sa. Mărimea $N_a(\delta)$ reprezintă procesul binomial, respectiv cel aproximativ, de tip Poisson (vezi aplicaţia anterioară), de numărare a sosirilor pe intervalul δ.

- **Probabilitatea ca** pe parcursul unui interval infinitezimal, δ, **să se încheie o servire**. Formula de calcul:

$$P[N_d(\delta)=1 \mid N(t)>0] = P[\tau\le\delta] = 1-e^{-\mu\cdot\delta} \cong \mu\cdot\delta + o(\delta) \qquad (12.11.4)$$

unde τ este o variabilă exponenţială care reprezintă timpul de servire al tranzacţiei cu care este ocupată unitatea de servire a sistemului şi $N_d(t)$ este procesul de numărare al cererilor satisfăcute în intervalul δ. Şi în acest caz, expresia pune în evidenţă faptul că probabilitate în cauză este semnificativă.

1. **Probabilitatea ca** pe parcursul unui interval infinitezimal, δ, **o cerere să sosească şi o cerere să fie servită**. Dat fiind independenţa proceselor de sosire şi servire, formula de calcul conduce la următorul rezultat:

$$P[N_a(\delta)=1, N_d(\delta)=1 \mid N(t)=k] = P[N_a(\delta)=1 \mid N(t)=k]\cdot P[N_d(\delta)=1 \mid N(t)=k] = \ldots \cong o(\delta) \qquad (12.11.5)$$

care specifică faptul că şi această probabilitate este nesemnificativă.

2. **Probabilitatea sosirii mai multor cereri** pe parcursul unui interval infinitezimal, δ. Descriind sosirile prin intermediul aceluiaşi proces Poisson, formula de calcul este, pentru $j\ge 2$ şi $j\le c-k$:

$$P[N_a(\delta)=j \mid N(t)=k] = \frac{[(c-k)\cdot\alpha\cdot\delta]^j}{j!}\cdot e^{-(c-k)\cdot\alpha\cdot\delta} = o(\delta) \qquad (12.11.6)$$

În concluzie, aceste probabilităţi sunt nesemnificative.

3. **Probabilitatea finalizării mai multor cereri** pe parcursul unui interval infinitezimal este δ. Acest eveniment este echivalent cu evenimentul:

$\{$*suma timpilor de servire a unui număr de* j *cereri să fie mai mică, cel mult egală cu durata* $\delta\}$, ceea ce permite scrierea următoarei egalităţi de probabilităţi pentru $j \geq 2$ şi $j \leq k$:

$$P\left[N_d(\delta) = j \mid N(t) = k\right] = P\left[\tau_1 + \tau_2 + \ldots + \tau_j \leq \delta\right] \quad (12.11.7)$$

Suma variabilelor aleatorii cuprinse în expresia celei de-a doua probabilităţi, dat fiind faptul că sunt i.i.d. după o lege de distribuţie exponenţială, reprezintă o variabilă de tip *j*-Erlang, notată, de exemplu, X, şi descrisă de funcţia de densitate de probabilitate:

$$f_X(x) = \mu \cdot e^{-\mu \cdot x} \cdot (\mu \cdot x)^{j-1} / (j-1)! \quad (12.11.8)$$

Prin urmare, probabilitatea căutată poate fi determinată calculând următoarea integrală (prin părţi):

$$P[X \leq \delta] = \int_0^{\delta} f_X(x)\,dx = e^{-\mu \cdot \delta} \cdot (\mu \cdot \delta)^j / j! \quad (12.11.9)$$

Rezultatul se poate obţine mai uşor dacă ţinem cont că numărul cererilor servite (deci, ieşite din sistem) pe parcursul unui interval, urmează un proces Poisson de rată μ. În acest caz:

$$P\left[N_d(\delta) = j \mid N(t) = k\right] = (\mu \cdot \delta)^j / j! \cdot e^{-\mu \cdot \delta} = o(\delta) \quad (12.11.10)$$

Conform rezultatului obţinut, şi probabilităţile acestor evenimente sunt neglijabile.

Bilanţul probabilităţilor fiind încheiat, se poate trage concluzia finală că procesul analizat aparţine clasei de procese de naştere şi moarte.

* * *

Concluzia din aplicaţia precedentă permite trasarea diagramei de tranziţii din figura 12.11.3.

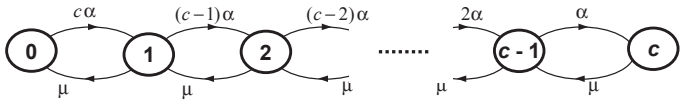


Figura 12.11.3: Diagrama ratelor de tranziţii – M/M/1/c

Aplicaţia 12.11.3

Scrieţi ecuaţiile de echilibru global şi deduceţi expresia de calcul a probabilităţilor de stare pentru procesul corespunzător diagramei din figura 12.11.3.

Răspuns: $p_k\big|_c = P[N = k] = p_0 \cdot c!/(c-k)! \cdot (\alpha/\mu)^k,\ 0 \leq k \leq c$ (12.11.11)

unde: $$p_0 = \left\{\sum_{j=0}^{c} (\alpha/\mu)^j \cdot c!/(c-j)!\right\}^{-1} \quad (12.11.12)$$

Notaţia N corespunde procesului aleatoriu care reprezintă numărul de cereri aflate la un moment dat în sistem. Indicele inferior, c, care apare în notaţia folosită pentru probabilitatea de stare, este opţional, putând fi omis pentru simplificarea notaţiei, atunci când nu sunt motive de confuzie, şi reprezintă numărul considerat al terminalelor (clienţi potenţiali).

* * *

Performanţa unui astfel de model se evaluează prin intermediul indicatorilor următori:

1) ***Traficul servit*** sau ***gradul de utilizare*** a unităţii de servire, g, reprezintă, ca frecvenţă relativă, procentul stabilit pe termen lung cât sistemul este ocupat, fiind, prin urmare, dat, din punct de vedere probabilistic, de relaţia:

$$g = 1 - p_0 \quad (12.11.13)$$

2) ***Productivitatea sistemului,*** γ**.** Sistemul având o singură unitate de servire, gradul de utilizare al acesteia reprezintă şi numărul mediu de clienţi aflaţi în ea. Prin urmare, aplicând formula Little la nivelul acestui subsistem (notat s în figura 12.11.2), obţinem relaţia:

$$g = \gamma \cdot E[\tau], \text{ unde } E[\tau] = 1/\mu \quad (12.11.14)$$

din care rezultă expresia căutată:

$$\gamma = \mu \cdot (1 - p_0) \quad (12.11.15)$$

De remarcat că, sistemul servind fără pierderi, rata de intrare a cererilor în sistem, λ, este egală cu rata de plecare, λ_d.

3) ***Timpul mediu petrecut de un client (cerere de serviciu) în sistem,*** $E[T]$**.** Stabilirea expresiei de calcul a acestei mărimi, numită şi **timp mediu de răspuns**, ţine cont de faptul că diferenţa medie dintre momentele iniţierii a două cereri succesive generate de aceeaşi sursă, care reprezintă ciclul mediu de activitate, corespunde sumei dintre mărimea căutată, $E[T]$, şi durata medie dintre momentul încheierii unei tranzacţii şi momentul iniţierii cererii următoare, adică $1/\alpha$. În consecinţă, fiecare sursă generează cereri cu o rată egală cu inversul sumei amintite mai sus ceea ce, la nivelul întregului ansamblu de surse, se traduce într-o rată medie de intrare a cererilor în sistem de expresie:

$$\lambda = \frac{c}{1/\alpha + E[T]} \quad (12.11.16)$$

Întrucât această rată este egală cu rata de plecare, λ_d, sistemul servind fără pierderi, rezultă că timpul mediu petrecut de un client în sistem are următoarea formulă de calcul:

$$E[T] = c/\gamma - 1/\alpha \quad (12.11.17)$$

4) ***Numărul mediu de clienţi (cereri) în sistem,*** $E[N]$**.** Se aplică, desigur, formula Little obţinându-se:

$$E[N] = c - \gamma/\alpha \quad (12.11.18)$$

Cunoscând acest număr putem stabili, calculând diferenţa $c - E[N]$, ***numărul mediu de terminale în repaus***.

5) ***Timpul mediu de aşteptare***, $E[W]$**.** La nivel de medii se poate scrie următoarea relaţie:

$$E[W] = E[T] - 1/\mu \quad (12.11.19)$$

6) ***Probabilitatea stării de blocare*** (ocupare, BUSY) a unui client, p_b. Din punct de vedere probabilistic, comportamentul unei surse poate fi descris printr-un lanţ Markov cu două stări, a cărei diagramă este reprezentată în figura 12.11.4. O astfel de modelare este posibilă întrucât proprietatea de ***lipsa a memoriei*** caracterizează timpii de *staţionare* în ambele stări în care se poate găsi o sursă.

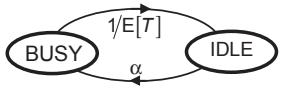


Figura 12.11.4: Modelarea activităţii unui client

Astfel, în cazul stării IDLE durata de staţionare are media $1/\alpha$ (conform ipotezei) şi urmează o distribuţie exponenţială, iar în cazul BUSY durata de staţionare are media $\mathrm{E}[T]$ şi corespunde unei sume (variabila T) în număr variabil de variabile aleatorii distribuite exponenţial de medie $1/\mu$. Prin urmare, rezolvând sistemul alcătuit din ecuaţia de echilibru şi condiţia de normare, se obţine relaţia:

$$p_b = P[\mathrm{BUSY}] = \frac{\mathrm{E}[T]}{\mathrm{E}[T] + 1/\alpha} \qquad (12.11.20)$$

Aplicaţia 12.11.4

Să se reprezinte grafic folosind MATLAB variaţia cu numărul de clienţi, c, a ratei de plecare, γ, şi a timpului de răspuns, $\mathrm{E}[T]$, raportat la durata de servire, $1/\mu$, care caracterizează modelul M/M/1/*c*.

Indicaţii: Se vor concepe fişiere funcţie, care vor primi ca argumente mărimile α şi μ.

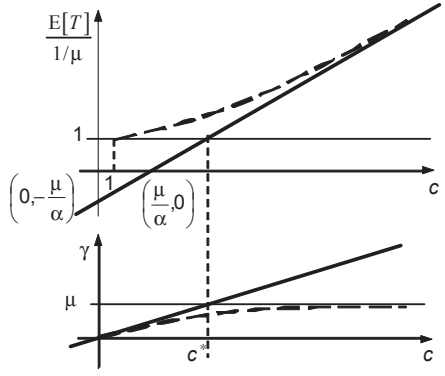


Figura 12.11.5: Variaţia cu numărul de clienţi potenţiali a întârzierii medii normate şi a ratei de plecare corespunzătoare modelului M/M/1/c

Corectitudinea curbelor trasate se poate stabili verificând dacă acestea sunt tangente la două asimptote, ce se construiesc luând în considerare cazurile extreme: c foarte mic şi c foarte mare. Astfel, în primul caz aşteptarea în sistem este un eveniment foarte rar, motiv pentru care se poate considera $\mathrm{E}[W] \cong 0$; această observaţie permite deducerea asimptotelor corespunzătoare:

$$\mathrm{E}[T] = 1/\mu - \mathrm{E}[W] \cong 1/\mu \text{ şi } \gamma = \lambda = c/(1/\alpha + 1/\mu)$$

Al doilea caz se referă la situaţia în care resursa sistemului serveşte aproape continuu, deci: $\gamma = \lambda \cong \mu$ şi $\mathrm{E}[T] = c/\gamma - 1/\alpha$.

Dreptele descrise prin ecuaţiile stabilite anterior sunt desenate în figura 12.11.5. Aceasta conţine, la modul calitativ, şi reprezentarea cu linie întreruptă a curbelor reale. Punctul de intersecţie a celor două perechi de asimptote, notat c^*, a cărei expresie de calcul rămâne în sarcina cititorului, se consideră a fi delimitatorul *zonei de congestie* a sistemului.

12.12 Modelul M/G/1

Modelul M/G/1 este un sistem cu aşteptare la care sosirile urmează un proces Poisson, iar timpii de servire $\tau_1, \tau_2, \cdots$ sunt variabile aleatorii independente şi identic distribuite după o lege specificată printr-o anumită densitate de probabilitate $f_\tau(x)$ şi având primele două momente finite.

În cazul modelelor de tip **M/M/...,** starea, ca informaţie pentru stabilirea probabilităţilor evenimentelor viitoare, este specificată doar de numărul de clienţi în sistem, $N(t)$. Acest lucru este posibil deoarece distribuţia exponenţială ce guvernează atât timpii de servire cât şi cei de sosire se caracterizează prin lipsa memoriei. Astfel, timpul rezidual rămas până la ieşirea dintr-o anumită stare, contorizat dintr-un moment t, ales pe parcursul staţionării în aceasta, urmează aceeaşi distribuţie cu cea a timpului de staţionare τ, contorizat din momentul intrării în respectiva stare. Exprimarea matematică a acestei proprietăţi este:

$$P[\tau > t + x \mid \tau > t] = P[\tau > x]$$

Însă, în cazul sistemului M/G/1, această proprietate nu mai este valabilă. Din acest motiv, **timpul rezidual**, τ_r, contorizat din momentul t, reprezintă a doua componentă care, alături de numărul de clienţi în sistem, specifică starea sistemului la momentul t. În consecinţă, probabilitatea de stare a unui sistem M/G/1 se exprimă, la modul general, sub forma: $P(N(t) = n, \tau_r(t) \le x)$.

Dar o asemenea exprimare a probabilităţii este dificil de folosit în calcule şi, de aceea, se preferă stabilirea altor relaţii mai uşor de utilizat şi care să ne permită a determina corect distribuţiile principalelor mărimi ce caracterizează un sistem M/G/1.

Ecuaţiile Pollaczek-Khinchin de transformare răspund acestui deziderat, oferind formule de calculul pentru următoarelor mărimi:

- $G_N(z)$, funcţia generatoare a probabilităţilor numărului de clienţi prezenţi în sistem :

$$G_N(z) = \frac{(1-\rho)\cdot(z-1)\cdot\hat{\tau}(\lambda\cdot(1-z))}{z-\hat{\tau}(\lambda\cdot(1-z))} \qquad (12.12.1)$$

- $\hat{T}(s)$, transformata Laplace a funcţiei densitate de probabilitate a timpului petrecut de un client în sistem :

$$\hat{T}(s) = \frac{(1-\rho)\cdot s\cdot\hat{\tau}(s)}{s-\lambda+\lambda\cdot\hat{\tau}(s)} \qquad (12.12.2)$$

- $\hat{W}(s)$, transformata Laplace a funcţiei densitate de probabilitate a timpului de aşteptare în sistem:

$$\hat{W}(s) = \frac{(1-\rho)\cdot s}{s-\lambda+\lambda\cdot\hat{\tau}(s)} \qquad (12.12.3)$$

în care: τ = variabila aleatorie ce reprezintă timpul de servire,

$\hat{\tau}(s)$ = transformata Laplace a densităţii de probabilitate a timpului de servire,

$\rho = -\lambda\cdot\hat{\tau}'(s)\big|_{s=0} = \lambda\cdot\mathrm{E}[\tau]$ este traficul oferit. (12.12.4)

Prin procedee analitice sau numerice, se aplică transformata inversă asupra acestor relaţii şi se obţin distribuţiile căutate:

$$P[N(t)=k],\ f_T(x) \text{ şi } f_w(x).$$

Relaţia (12.12.1) se deduce [Doy89] transformând analiza din domeniul procesului original în cel al unui proces asociat (**inclus** sau **încuibărit**, numit ***embedded*** în terminologia engleză) care, în cazul modelului M/G/1, reprezintă numărul clienţilor din sistem la momentele t de finalizare ale servirilor (când timpul rezidual $\tau_r(t)$ este nul). Pentru procesul asociat în discuţie, tranziţiile între stări sunt descrise prin intermediul diagramelor din figura 12.12.1.

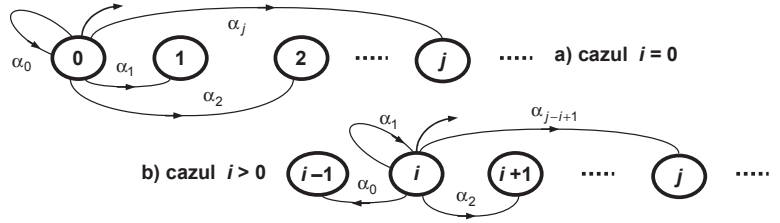


Figura 12.12.1: Diagrama de tranziţii ale procesului asociat modelului M/G/1

Valorile α_k ale probabilităţilor cu care au loc respectivele tranziţii între stări se determină prin intermediul relaţiei:

$$\alpha_k = P[A(\tau)=k] = \int_{S_\tau} P[A(x)=k\,|\,\tau=x]\cdot f_\tau(x)\,\mathrm{d}x \qquad (12.12.4)$$

unde $A(\tau)$ este procesul de numărare a sosirilor de la o servire la alta, a cărui probabilitate elementară (de ordinul 1) se determină luând în considerare toate realizările posibile, cuprinse în domeniul S_τ, ale timpului de servire, τ, şi aplicând teorema probabilităţii totale.

Pe baza diagramelor din figura 12.12.1, se deduce expresia generală a ecuaţiilor de echilibru global:

$$p_j = p_0\cdot\alpha_j + \sum_{i=1}^{j+1} p_i\cdot\alpha_{j-i+1} \qquad (12.12.5)$$

care, dat fiind imposibilitatea stabilirii unei relaţii de recurenţă între probabilităţi, se rezolvă utilizând transformata Z.

Pentru aceasta, ansamblul ecuaţiilor descrise generic prin intermediul relaţiei (12.12.5) se amplifică cu z^j şi apoi se adună membru cu membru obţinându-se egalitatea:

$$\underbrace{\sum_{j=0}^{\infty} p_j\cdot z^j}_{G(z)} = p_0\cdot\underbrace{\sum_{j=0}^{\infty}\alpha_j\cdot z^j}_{G_A(z)} + \sum_{j=0}^{\infty}\sum_{i=1}^{j+1} p_i\cdot\alpha_{j-i+1}\cdot z^j \qquad (12.12.6)$$

În cadrul acestei egalităţi, membrul stâng reprezintă funcţia generatoare, $G(z)$, a probabilităţilor corespunzătoare numărului de clienţi din sistem, în momentul finalizării unei serviri, iar membrul drept cuprinde doi termeni dintre care în primul găsim funcţia generatoare, $G_A(z)$, a probabilităţilor corespunzătoare numărului de sosiri pe durata unei serviri. Pentru această funcţie generatoare se poate scrie următorul şir de egalităţi:

$$G_A(z) = \sum_{j=0}^{\infty}\alpha^j\cdot z^j = \sum_{j=0}^{\infty} z^j\int_0^{\infty}\frac{(\lambda x)^j}{j!}\cdot\mathrm{e}^{-\lambda x}\cdot f_\tau(x)\,\mathrm{d}x$$

$$= \int_0^{\infty}\sum_{j=0}^{\infty}\frac{(\lambda z x)^j}{j!}\cdot\mathrm{e}^{-\lambda x}\cdot f_\tau(x)\,\mathrm{d}x = \int_0^{\infty} f_\tau(x)\cdot\mathrm{e}^{-\lambda(1-z)x}\,\mathrm{d}x$$

în care, ultima expresie se poate pune sub forma:

$$G_A(z) = \hat{\tau}(\lambda\cdot(1-z)) \qquad (12.12.7)$$

ceea ce înseamnă că funcţia generatoare $G_A(z)$ are expresia dată de transformata Laplace a densităţii de probabilitate a timpului de servire $\hat{\tau}(s)$, cu argument modificat $s \leftarrow \lambda\cdot(1-z)$.

În privinţa celui de al doilea termen din membrul drept al relaţiei (12.12.5), el poate fi pus sub forma: $\frac{1}{z}\cdot\sum_{i=1}^{\infty} p_i\cdot z^i\cdot\sum_{m=0}^{\infty}\alpha_m\cdot z^m$, care s-a obţinut prin inversarea sumelor şi adoptând substituţia $j = m+i-1$.

În această formă apare din nou necunoscuta $G(z)$, dat fiind faptul că:

$$\sum_{i=1}^{\infty} p_i \cdot z^i = G(z) - p_0 .$$

Având la dispoziţie aceste rezultate intermediare, relaţia (12.12.6) poate fi rescrisă în maniera:

$$G(z) = p_0 \cdot G_A(z) + 1/z \cdot \big(G(z) - p_0\big) \cdot G_A(z)$$

din care se obţine:

$$G(z) = p_0 \cdot \frac{(z-1)\cdot G_A(z)}{z - G_A(z)} \qquad (12.12.8)$$

Mărimea p_0 se determină aplicând relaţiei (12.12.8) condiţia de normare, adică particularizând-o pentru $z = 1$; astfel, după înlăturarea nedeterminării $0/0$ utilizând regula lui l'Hôpital, se obţine egalitatea:

$$p_0 = \frac{1 - G_A'(z)}{(z-1)\cdot G_A'(z) + G_A(z)} \qquad (12.12.9)$$

în care: $G_A(z)\big|_{z=1} = 1$, şi

$$G_A'(z)\big|_{z=1} = \mathrm{E}[A] = \int_0^{\infty} \underbrace{\mathrm{E}\big[A(x)\big|\tau = x\big]}_{\lambda \cdot x} \cdot f_\tau(x)\,\mathrm{d}x = \underbrace{\lambda \cdot \mathrm{E}[\tau]}_{\rho}.$$

Cu aceste relaţii, probabilitatea p_0, ca sistemul să fie vid, este dat de relaţia $p_0 = 1 - \rho$ şi, în final, utilizând şi substituţia permisă de relaţia (12.12.7) se obţine relaţia (12.12.1).

Desigur că pe parcursul demonstraţiei anterioare, $G(z)$ a reprezentat funcţia generatoare corespunzătoare distribuţiei staţionare de ordinul unu a procesului corespunzător *numărului de clienţi prezenţi în sistem în momentele servirii*, ND_j. Însă această distribuţie corespunde şi procesului $N(t)$, număr de clienţi în sistem, deoarece sunt adevărate următoarele afirmaţii:

1. *distribuţia numărului de clienţi găsiţi de un client la sosire*, NA_n, este identică cu distribuţia stărilor procesului $N(t)$, atunci când procesul sosirilor este independent de stările sistemului.
2. pentru sistemele în care sosirile şi plecările sunt individuale, distribuţia NA_n, număr de clienţi găsiţi de un client la sosire, este aceeaşi cu distribuţia ND_n, număr de clienţi rămaşi la plecarea unui client.

Valabilitatea ambelor afirmaţii se poate verifica odată cu rezolvarea aplicaţiilor 12.17.7 şi 12.17.8.

Cea de a doua ecuaţie Pollaczek-Khinchin, (12.12.2), se determină plecând de la faptul că numărul ND_j de clienţi rămaşi în sistem după ieşirea unui client oarecare j este egal cu numărul de clienţi intraţi în sistem pe durata

aleatorie, T, petrecută în sistem de respectivul client. În consecinţă, probabilitatea ca în urma ieşirii clientului j în sistem să rămână k clienţi are o expresie asemănătoare cu relaţia (12.12.4) şi anume:

$$P\big[A(T_j) = k\big] = \int_{S_T} P\big[A(t) = k \,\big|\, T = t\big] \cdot f_T(t)\,\mathrm{d}t \qquad (12.12.10)$$

unde S_T este spaţiul realizărilor variabilei aleatorii T.

Prin urmare, conform raţionamentului folosit la deducerea relaţiei (12.12.7), se obţine următoarea expresie a funcţiei generatoare corespunzătoare distribuţiei staţionare de ordinul unu a procesului ND_n:

$$G_{ND}(z) = \hat{T}\big(\lambda \cdot (1-z)\big) \qquad (12.12.11)$$

unde $\hat{T}(s)$ este transformata Laplace a densităţii de probabilitate a variabilei aleatorii T. Conform discuţiei legate de stabilirea relaţiei (12.12.1), $N(t)$ şi ND_n au aceeaşi distribuţie ceea ce permite, utilizând relaţiile (12.12.1) şi (12.12.11), împreună cu substituţia $s = \lambda \cdot (1-z)$, să se ajungă la relaţia (12.12.2).

Ultima ecuaţie Pollaczek-Khinchin, (12.12.3), se deduce plecând de la relaţia $T = W + \tau$, în care variabilele W şi τ sunt independente. Din acest motiv:

$$\hat{T}(s) = \hat{W}(s) \cdot \hat{\tau}(s) \qquad (12.12.12)$$

ceea ce conduce automat la relaţia (12.12.3).

Aplicaţia 12.12.1

Folosiţi ecuaţiile Pollaczek-Khinchin pentru a determina distribuţiile mărimilor următoare:

i) $N(t)$, număr de clienţi în sistem,
ii) T, timp petrecut în sistem,
iii) W, timp de aşteptare care caracterizează modelul M/M/1.

Rezolvare: Corespunzător modelului M/M/1, timpul de servire, τ, urmează o distribuţie exponenţială, $f_\tau(t) = \mu \cdot \exp(-\mu \cdot t)$, a cărei transformată Laplace este:

$$\hat{\tau}(s) = \int_0^{\infty} \mathrm{e}^{-s\cdot t} \cdot f_\tau(t)\,\mathrm{d}t = \mu \cdot \int_0^{\infty} \mathrm{e}^{-(s+\mu)\cdot t}\,\mathrm{d}t = \frac{\mu}{s+\mu} \text{ şi deci:}$$

i) $$G_N(z) = \frac{(1-\rho)(z-1)\cdot \dfrac{\mu}{\lambda(1-z)+\mu}}{z - \dfrac{\mu}{\lambda(1-z)+\mu}} = \frac{1-\rho}{1-\rho\cdot z} \text{ unde } \rho = \lambda/\mu .$$

Dezvoltând $G_N(z)$ în serie de puteri rezultă următoarea expresie de calcul a probabilităţilor de stare:

$$P[N(t) = k] = (1-\rho)\cdot\rho^k,\ k = 0,1,2,\cdots$$

expresie care este identică cu cea obţinută în cadrul analizei modelului M/M/1.

ii) $\hat{T}(s)=\dfrac{(1-\rho)\cdot s\mu}{(s+\mu)\cdot(s-\lambda)+\lambda\mu}=\dfrac{(1-\rho)\mu}{s-(\lambda-\mu)}$. Cunoscând că $1/s \overset{L}{\leftrightarrow} \sigma(x)$, unde $\sigma(x)$ este funcţia treaptă unitate, şi aplicând proprietatea translatării în spaţiul s, $a(x)\cdot e^{+s_0\cdot x} \overset{L}{\leftrightarrow} A\cdot(s-s_0)$, se obţine:

$$f_T(x)=\mu(1-\rho)\cdot\exp\left[-\mu(1-\rho)x\right], \text{ pentru } x>0.$$

iii) $W(s)=\dfrac{(1-\rho)\cdot s}{s-\lambda+\lambda\mu/(s+\mu)}=(1-\rho)\cdot\dfrac{s+\mu}{s+\mu-\lambda}=(1-\rho)\cdot\left(1-\dfrac{\lambda}{s+\mu-\lambda}\right)$

Cunoscând că $\delta(x) \overset{L}{\leftrightarrow} 1$, unde $\delta(x)$ este impulsul Dirac (delta), se obţine:

$$f_W(x)=(1-\rho)\cdot\delta(x)+\lambda(1-\rho)\cdot\exp\left[-\mu(1-\rho)x\right],\ x\geq 0.$$

Conform acestei expresii, variabila aleatorie W este mixtă, luând:

- valoarea 0 cu probabilitate $1-\rho$,
- restul valorilor după o funcţie de densitate de probabilitate (al doilea termen din expresie).

* * *

Ţinând cont de proprietăţile transformatei Laplace, media timpului de aşteptare, corespunzătoare modelului M/G/1, este dată de relaţia:

$$\mathrm{E}[W]=-\mathrm{d}\hat{W}(s)/\mathrm{d}s\big|_{s=0}$$

care conduce, prin calcule care implică şi regula lui l'Hospital, la expresia:

$$\mathrm{E}[W]=\frac{\lambda}{2(1-\rho)}\cdot\mathrm{E}\left[\tau^2\right]=\frac{\rho\cdot(1+c_\tau^2)}{2(1-\rho)}\cdot\mathrm{E}[\tau] \qquad (12.12.13)$$

în care $c_\tau=\mathrm{STD}[\tau]/\mathrm{E}[\tau]$ este ***coeficientul de variaţie*** a timpului de servire.

Cu ajutorul mărimii $\mathrm{E}[W]$ se poate determina întârzierea medie introdusă de sistem: $\mathrm{E}[T]=\mathrm{E}[W]+\mathrm{E}[\tau]$, şi cunoscând $\mathrm{E}[T]$ se poate stabili numărul mediu de clienţi în sistem, şi anume: $\mathrm{E}[N]=\lambda\cdot\mathrm{E}[T]$

Aplicaţia 12.12.2

Fie un multiplexor statistic de celule (blocuri informaţionale de aceeaşi lungime) care provin de la trei categorii de surse. Să se determine performanţa sistemului ştiind că:

- sursele din cele trei categorii generează celule după o distribuţie Poisson de medie: $\lambda_1=80$ celule/sec, $\lambda_2=400$ celule/sec şi $\lambda_3=2.500$ celule/sec;
- numărul surselor active din fiecare categorie este: $n_1=125$, $n_2=50$ şi $n_3=20$;
- durata de servire a unei celule este constantă şi are valoarea $\tau=10\,\mu$sec.

Rezolvare: Cu datele precizate se calculează următoarele:

- rata totală de intrare în sistem: $\lambda=n_1\lambda_1+n_2\lambda_2+n_3\lambda_3=8\cdot10^4$ celule/sec;
- gradul de utilizare a canalului de transmisiuni: $\rho=\lambda\,\mathrm{E}[\tau]=\lambda\cdot\tau=0{,}8\,[\mathrm{E}]$;
- coeficientul de variaţie a timpului de servire: $c_\tau^2=\mathrm{VAR}[\tau]/\mathrm{E}^2[\tau]=0$;

şi de aici rezultă următoarele valori pentru indicatorii de performanţă:

$$\mathrm{E}[W]=\tau\cdot\rho/\left[2(1-\rho)\right]=20\,\mu\text{sec}$$

* * *

Aplicaţia 12.12.3

Să se determine performanţele unui sistem cu capacitatea infinită de stocare, la care servirea se desfăşoară într-o singur unitatea şi presupune parcurgerea a două faze, în conformitate cu structura Cox, unde $\alpha_1=1$ şi $\alpha_2=p$, iar sosirile urmează un proces Poisson de rată λ.

* * *

Figura 12.12.2(a) ilustrează faptul că pentru un şir M/G/1 numărul mediu de clienţi prezenţi (sarcini de tratat) creşte liniar cu pătratul coeficientului de variaţie şi că rata de creştere este foarte sensibilă la mărimea ρ a utilizării serverului. De asemenea, din imaginea (b) a figurii se poate observa că, dacă utilizarea tinde spre 1, apare o creştere dramatică a numărului de clienţi şi aceasta se petrece cu atât mai repede, cu cât coeficientul de variaţie este mai mare. Înseamnă că trebuie impusă o limită a valorii utilizării, pentru a nu produce congestionarea sistemului de servire. Acest fenomen apare de fapt şi în cazul şirului M/M/1, unde curba de variaţie a numărului de clienţi prezenţi în şir are exact aceeaşi formă, prezentând o creştere asimptotică la $\rho=1$.

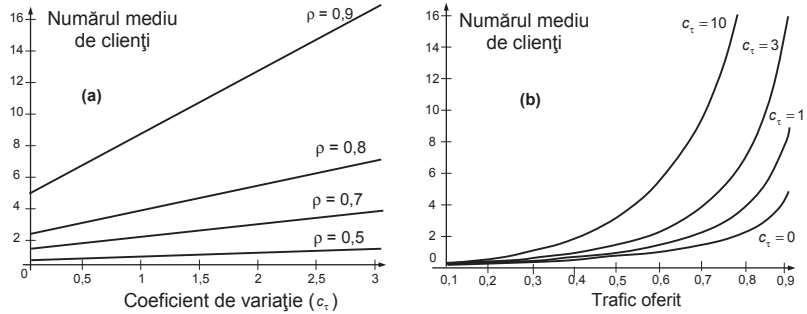


Figura 12.12.2: Variaţia numărului de clienţi prezenţi în şirul M/G/1

În cazul timpilor de servire având distribuţii cu cozi semnificative (***heavy-tailed distribution***), de genul $P[T>x]\sim L(x)\cdot x^{-\nu}$, cu $L(x)$ având variaţii lente, ($\lim_{t\to\infty} L(tx)/L(t)=1$, pentru $\forall\, x>0$), următoarea relaţie aproximativă [BBQH03]

se verifică privind timpul de aşteptare: $P[W > x] \sim C \cdot x^{1-\nu}$

Drept urmare, distribuţia timpului de aşteptare, ca şi cea a timpului de tranzit, sunt cu un grad mai semnificative (***one degree heavier***) decât distribuţia timpului de servire.

12.13 Modelul GI/M/1

În cazul şirurilor GI/M/1, unde **GI** (***G**eneral **I**ndependent*) specifică faptul că timpii dintre două sosiri succesive, $\{A_i\}$, reprezintă o secvenţă de variabile aleatorii i.i.d. după o lege generală, se foloseşte parametrul σ dat de relaţia:

$$\sigma = \hat{A}(\mu - \mu\sigma) \qquad (12.13.1)$$

unde $\hat{A}(s)$ este transformata Laplace a timpului dintre două sosiri succesive.

În consecinţă, considerând gradul de utilizare al unităţii de servire, $\rho = (\mathrm{E}[A]\cdot\mu)^{-1}$, probabilităţile stărilor şirului sunt definite cu relaţiile următoare:

$$p_k = \rho\cdot(1-\sigma)\cdot\sigma^{k-1}, \text{ pentru } k>0, \text{ şi } p_0 = 1-\rho \qquad (12.13.2)$$

iar funcţia de distribuţie a timpului de aşteptare este dată de expresia:

$$F_w(x) = \begin{cases} 1-\sigma & \text{pentru } x = 0 \\ 1-\sigma\cdot\exp[-\mu(1-\sigma)x] & \text{pentru } x > 0 \end{cases} \qquad (12.13.3)$$

Ca urmare, **indicatorii de performanţă** ai şirului sunt calculaţi cu relaţiile:

- **numărul mediu de clienţi în sistem** şi **varianţa** lui:

$$\mathrm{E}[N] = \rho/(1-\sigma) \text{ şi } \mathrm{VAR}[N] = \rho\cdot(1+\sigma-\rho)\cdot(1-\rho)^{-2} \qquad (12.13.4)$$

- **timpul mediu de răspuns** (de tranzit al clienţilor):

$$\mathrm{E}[T] = \{\mu\cdot(1-\sigma)\}^{-1} \qquad (12.13.5)$$

- **lungimea medie** şi **varianţa şirului de aşteptare**:

$$\mathrm{E}[N_q] = \rho\cdot\sigma/(1-\sigma) \text{ şi } \mathrm{VAR}[N_q] = \rho\cdot\sigma\cdot(1+\sigma-\sigma\rho)\cdot(1-\rho)^{-2} \qquad (12.13.6)$$

- **timpul mediu de aşteptare**:

$$\mathrm{E}[W] = \sigma\cdot[\mu\cdot(1-\sigma)]^{-1} \qquad (12.13.7)$$

Comportarea şirurilor M/G/1 şi GI/M/1 este foarte asemănătoare, mai ales dacă $c_X^2 \le 1$, ceea ce se poate observa în figura 12.13.1, unde sunt trasate alăturat variaţiile numărului mediu de clienţi, pentru aceeaşi valoarea a coeficientului c_X^2, variabila X reprezentând, funcţie de modelul considerat, timpul de servire, respectiv timpul dintre două sosiri succesive.

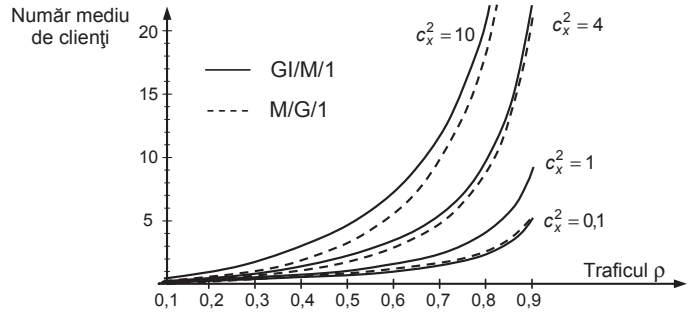


Figura 12.13.1: Numărul mediu de clienţi în şirurile M/G/1 şi GI/M/1

Aplicaţia 12.13.1

Folosind instrumentarul de calcul propriu modelului GI/M/1, stabiliţi formulele de evaluare a indicatorilor de performanţă ce caracterizează modelul: i) M/M/1; ii) E_2/M/1; iii) H_2/M/1.

Soluţie; ii) $\sigma = \rho + 0{,}5 - \sqrt{\rho - 0{,}25}$ ş.a.m.d.

12.15 Integrarea serviciilor

În practică, există o multitudine de posibilităţi de a multiplexa (combina) tipuri diferite de trafic. La nivelul "legătură de date" (*Link Layer* în modelul OSI) componentele unui trafic *eterogen* pot fi:

- distribuite la canale distincte, de lărgimi de bandă corespunzătoare;
- înmagazinate în sisteme cu aşteptare cu una sau mai multe resurse (canale) pentru a urma o anumită disciplină în servirea lor (FIFO, LIFO, cu prioritate etc.);
- repartizate aceloraşi resurse, prin partajarea dinamică a acestora.

Subcapitolul 12.12.2 a tratat un mod particular de integrare a serviciilor, folosit în cazul reţelelor ce folosesc doar comutaţia de pachete, mecanismul considerat fiind cel al acordării de priorităţi. Integrarea se realizează şi în cazul sistemelor în care coexistă cele două tipuri de comutaţie, de pachete şi de circuite, modul concret de implementare şi performanţele caracteristice fiind prezentate în cele ce urmează prin intermediul câtorva exemple.

Figura 12.15.1 prezintă un exemplu de combinare (multiplexare), într-un cadru **TDM** (***T**ime **D**ivision **M**ultiplex*) aparţinând unei căi de comunicaţie, a unui număr de M intrări destinate comutaţiei de circuite şi a unei intrări care deserveşte un sistem de servire cu aşteptare, alocat transferului de pachete.

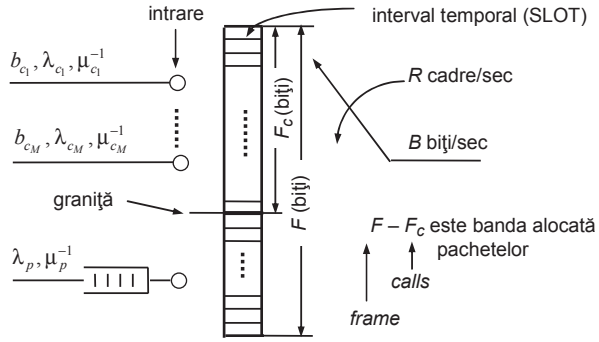


Figura 12.15.1: Exemplu de formare a unui cadru TDM

Traficul destinat comutaţiei de circuite este caracterizat de:

- lărgimea de bandă b_{c_i}, $i = \overline{1,M}$;
- timpul de servire (ocupare), de exemplu distribuit exponenţial, de medie $1/\mu_{c_i}$, $i = \overline{1,M}$;
- rata de sosire λ_{c_i}, corespunzătoare de exemplu unei distribuţii Poisson, $i = \overline{1,M}$.

După cum se observă, capacitatea, B, a căii de comunicaţie (de ieşire) este distribuită utilizatorilor de natură diferită, cărora li se alocă un număr corespunzător de intervale temporale per cadru, în funcţie de lărgimea de bandă cerută.

În practică se aplică diverse metode de integrare a acestor două tipuri de servicii (transmisii telefonice şi de date), a căror modelare matematică, [Schw96], în vederea evaluării performanţelor şi stabilirii unor relaţii de optimizare a proiectării, iau în considerare diverse ipoteze, în general, de genul:

- intrările alocate traficului telefonic au caracteristici comune: b_c, λ_c şi μ_c^{-1};
- apelurile telefonice folosesc, de la un cadru la altul, acelaşi canal pe toată durata lor de servire;
- pachetele de date sunt întârziate până ce apar canale disponibile pentru transmisia lor conform disciplinei aleasă;
- fiecare utilizator cere o bandă de 1 interval temporal per cadru;
- cadrul este alcătuit din s intervale temporale;
- procesele, precum transmiterea unui pachet sau desfăşurarea unui apel, se desfăşoară în mod continuu. Această ultimă ipoteză este valabilă cât timp durata unui cadru este neglijabilă faţă de timpul de servire. În acest caz, natura discretă a unui cadru TDM poate fi neglijată, sistemul putând fi analizat cu

ajutorul lanţurilor Markov continue în timp. De exemplu, considerând cadrul MIC primar, caracterizat prin: durată ($125\ \mu\text{sec}$), număr de intervale temporale (32) şi dimensiune a fiecărui interval temporal (8 biţi), atunci înseamnă că, dacă pentru transmisia unui pachet de 1200 biţi se foloseşte un singur interval temporal pe cadru, sunt necesare 150 cadre, ceea ce corespunde unui timp de servire de $16,75$ msec. Această valoare este mult mai mare decât durata unui cadru, făcând astfel să pară că pachetul este transmis în mod continuu pe un canal "izolat", de capacitate 64 kbiţi/sec.

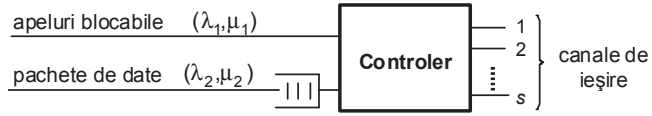


Figura 12.15.2: Model de principiu al unui sistem cu integrarea serviciilor

Luând în considerare ipotezele susmenţionate, schema model (de principiu) a unui sistem cu integrarea serviciilor de telefonie şi transmisiuni de date poate fi structurată în maniera prezentată în figura 12.15.2. Clienţii sistemului sunt repartizaţi controler-ului, fiind împărţiţi în două categorii:

- clasa (tipul) 1 - apeluri telefonice,
- clasa (tipul) 2 - pachete de date.

În funcţionarea controler-ului se poate opta pentru una din următoarele strategii, care însă nu epuizează toate soluţiile posibile:

- disciplina FIFO, LIFO etc., indiferent de tipul clientului;
- prioritate acordată apelurilor telefonice;
- specificarea unei clase de pachete a căror prioritate este superioară chiar şi apelurilor telefonice;
- specificarea unui număr s_1 de canale din cele s, care sunt alocate apelurilor telefonice, restul de $s - s_1$, inclusiv canalele rezervate serviciului telefonic, dar neutilizate temporar, fiind destinate transmisiei de pachete.

12.15.1. Integrarea utilizând disciplina FIFO

Modul de analiză a acestui gen de integrare a serviciilor este prezentat în continuare în cazul în care sistemul dispune de un singur canal ($s = 1$) şi de un şir infinit, servit fără prioritate, după o disciplină FIFO. Sunt considerate următoarele ipoteze:

- sosirile sunt de tip Poisson, atât ale apelurilor, cât şi ale pachetelor,
- timpii de servire sunt variabile aleatorii independente, identic distribuite exponenţial, de medie $1/\mu_1$ pentru apeluri şi $1/\mu_2$ pentru pachete.

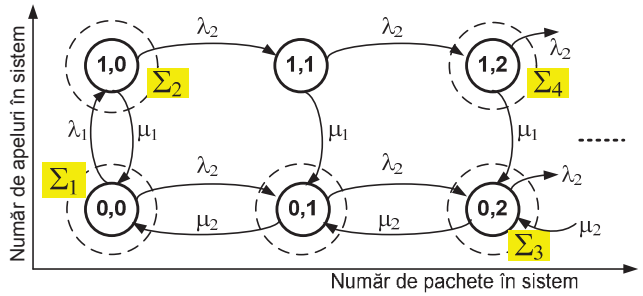


Figura 12.15.3: Diagrama de stări şi tranziţii în integrarea cu disciplină FIFO

În consecinţă, procesul evoluează conform diagramei de stări şi tranziţii prezentată în figura 12.15.3. Ecuaţiile globale de echilibru sunt :

- ecuaţia I, din Σ_1 : $(\lambda_1+\lambda_2)\cdot p_{0,0}=\mu_1\cdot p_{1,0}+\mu_2\cdot p_{0,1}$;
- ecuaţia II, din Σ_2 : $(\lambda_2+\mu_1)\cdot p_{1,0}=\lambda_1\cdot p_{0,0}$;
- ecuaţia III, din Σ_3, pentru $j\geq 1$:

$$(\lambda_2+\mu_2)\cdot p_{0,j}=\lambda_2\cdot p_{0,j-1}+\mu_2\cdot p_{0,j+1}+\mu_1\cdot p_{1,j};$$

- ecuaţia IV, din Σ_4, pentru $j\geq 1$: $(\lambda_2+\mu_1)\cdot p_{1,j}=\lambda_2\cdot p_{1,j-1}$.

Din ecuaţia II obţinem: $p_{1,0}=\lambda_1/(\mu_1+\lambda_2)\,p_{0,0}$, iar din ecuaţia IV rezultă relaţia de recurenţă:

$$p_{1,j}=\frac{\lambda_2}{\mu_1+\lambda_2}p_{1,j-1}=\left(\frac{\alpha\rho_2}{1+\alpha\rho_2}\right)^j\cdot\left(\frac{\rho_1}{1+\alpha\rho_2}\right)p_{0,0} \qquad (12.15.1)$$

în care: $\rho_1=\lambda_1/\mu_1$; $\rho_2=\lambda_2/\mu_2$; $\alpha=\mu_2/\mu_1$.

Deoarece pentru probabilităţile $p_{0,j}$, $j\geq 1$, ecuaţia III nu generează o relaţie de recurenţă, determinarea acestei probabilităţi se face utilizând transformata **Z** (funcţia generatoare). Astfel, multiplicând ecuaţiile III cu z^j şi sumându-le după $j\geq 1$ se obţine în final :

$$G_0(z)=p_{0,0}\cdot\left[1+\frac{\rho_1\rho_2\cdot z}{1-\alpha\rho_2\cdot(z-1)}\right]\cdot\frac{1}{1-\rho_2\cdot z} \qquad (12.15.2)$$

unde $G_0(z)$ este funcţia generatoare corespunzătoare probabilităţii $p_{0,j}$, adică:

$$G_0(z)=\sum_{j=0}^{\infty}p_{0,j}\,z^j$$

Probabilitatea $p_{0,0}$ se determină prin intermediul ecuaţiei de normare:

$$\sum_{i,j}p_{i,j}=\sum_{j=0}^{\infty}p_{0,j}+\sum_{j=0}^{\infty}p_{1,j}=1$$

în care: $\sum_{j=0}^{\infty}p_{0,j}=G_0(1)$ şi $\sum_{j=0}^{\infty}p_{1,j}=\rho_1\cdot p_{0,0}$. Se obţine astfel că:

$$p_{0,0}=(1-\rho_2)/(1+\rho_1) \qquad (12.15.3)$$

Evaluarea performanţei se realizează prin următorii indicatori:

❖ ***Probabilitatea de blocare*** (apeluri):

$$p_b=1-p_{0,0}=\underbrace{\frac{\rho_1}{1+\rho_1}}_{\substack{\text{blocare proprie}\\ \text{tip M/M/1/}\infty/1}}+\underbrace{\frac{\rho_2}{1+\rho_1}}_{\substack{\text{dispută apeluri/pachete}\\ \text{în obţinerea resursei}}} \qquad (12.15.4)$$

❖ ***Media numărului de pachete în sistem:***

$$\mathrm{E}[N_2]=\sum_{i,j}j\cdot p_{i,j}=\sum_{j=0}^{\infty}j\cdot p_{0,j}+\sum_{j=0}^{\infty}j\cdot p_{1,j}$$
$$=\frac{\mathrm{d}\,G_0(z)}{\mathrm{d}\,z}\bigg|_{z=1}+\frac{\alpha\rho_1\rho_2(1-\rho_2)}{1+\rho_2}$$

Deoarece: $\left.\frac{\mathrm{d}\,G_0(z)}{\mathrm{d}\,z}\right|_{z=1}=\frac{\rho_2}{1-\rho_2}+\frac{\alpha\rho_1\rho_2^2}{1+\rho_1}$, se obţine, în final:

$$\mathrm{E}[N_2]=\underbrace{\frac{\rho_2}{1-\rho_2}}_{\text{sistem M/M/1}}+\underbrace{\frac{\alpha\rho_1\rho_2}{1+\rho_1}}_{\substack{\text{dispută}\\ \text{pachete/apeluri}}} \qquad (12.15.5)$$

Ultima relaţie pune în evidenţă condiţiile în care sistemul atinge echilibrul statistic şi anume:

- ρ_1 este indiferent, servirea apelurilor făcându-se cu blocare,
- $\rho_2<1$, înmagazinarea pachetelor făcându-se într-o coadă infinită.

❖ ***Media timpului total petrecut de pachete în sistem:*** această mărime, raportată la durata medie de servire a unui pachet, se obţine cu ajutorul formulei lui Little, scrisă din punct de vedere al pachetelor ($\lambda_2\cdot\mathrm{E}[T_2]=\mathrm{E}[N_2]$) şi folosind mărimea anterioară, evaluată prin expresia (12.15.5):

$$\frac{\mathrm{E}[T_2]}{1/\mu_2}=\frac{\mathrm{E}[N_2]}{\rho_2}=\frac{1}{1-\rho_2}+\frac{\alpha\rho_1}{1+\rho_1} \qquad (12.15.6)$$

Pentru a interpreta relaţiile de calcul al performanţei, stabilite mai sus, se consideră un exemplu particular pentru care: $1/\mu_1=100\,\mathrm{sec}$, $1/\mu_2=10\,\mathrm{msec}$,

($\alpha = 10^4$), $\rho_1 = 0{,}1\text{E}$ şi $\rho_2 = 0{,}4\text{E}$. În aceste condiţii, obţinem:

- $p_b = 0{,}45$, adică o valoare mai mare decât cea care s-ar obţine dacă unicul canal ar fi destinat numai apelurilor, adică $p_b\big|_{\text{M/M/1/}\infty/1} \cong 0{,}1$;
- $\mu_2 \cdot \text{E}[T_2] = 1{,}67 + 909$, o valoarea cu mult mai mare decât cea care corespunde situaţiei de alocare a canalului numai apelurilor telefonice, $\mu_2 \cdot \text{E}[T_2]\big|_{\text{M/M/1}} = 1{,}67$.

Ambele rezultate evidenţiază faptul că într-un sistem cu un singur canal *interferenţa* între cele două tipuri de trafic este accentuată, performanţa fiind puternic afectată.

Pentru asigurarea unor timpi de întârziere şi probabilităţi de blocare acceptabile, în practică se utilizează mai multe canale ($s >> 1$). În acest caz, analiza sistemului generează un volum mare de calcule şi conduce la obţinerea unor expresii de dimensiuni considerabile, dificil de interpretat.

Din acest motiv, pentru astfel de situaţii s-au identificat relaţii aproximative de evaluare a indicatorilor de performanţă. Astfel, în cazul probabilităţii de blocare s-a observat că aceasta este relativ independentă de parametrul $\alpha = \mu_2/\mu_1$ pe o plajă largă de variaţie a mărimilor s, ρ_1 şi ρ_2. În consecinţă, relaţia de calcul al probabilităţii de blocare p_b, determinată în cazul $s = 1$, poate fi extinsă, ca aproximativă, pentru cazul $s \neq 1$, având forma:

$$p_b = \frac{\rho \cdot B(\rho, s-1)}{s - \rho_2 + \rho \cdot B(\rho, s-1)}, \text{ cu } \rho = \rho_1 + \rho_2 \qquad (12.15.7)$$

sau, exprimată în termeni ai relaţiei de recurenţă proprii formulei Erlang-B:

$$B(\rho, s) = \frac{\rho \cdot B(\rho, s-1)}{s + \rho \cdot B(\rho, s-1)} \qquad (12.15.8)$$

Aplicaţia 12.15.1

Verificaţi că pentru $s = 1$ se obţine expresia exactă, determinată anterior.

Observaţie: Dacă la intrare în sistem dispar pachetele se obţine expresia de recurenţă a formulei Erlang-B.

12.15.2. Integrarea cu prioritate absolută ($s = 1$)

Integrarea cu prioritate absolută este strategia în care controlerul oferă prioritate clienţilor de tipul 1 (apeluri), eliberând canalele de clienţii de tip 2 (pachete).

Luând în considerare ipotezele legate de procesele de sosire şi servire asumate în cazul anterior, al integrării fără prioritate, diagrama de tranziţii corespunzătoare acestei situaţii se prezintă ca în figura 12.15.4.

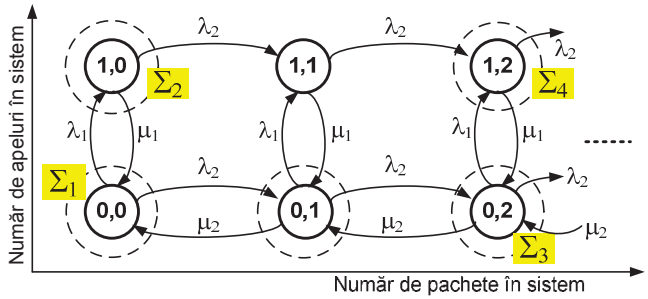


Figura 12.15.4: Diagrama de stări şi tranziţii în integrarea cu prioritate absolută

Aplicaţia 12.15.2

Scrieţi ecuaţiile de echilibru global pentru modelul integrării cu prioritate absolută, folosind diagrama de stări şi tranziţii din figura 12.15.4.

* * *

Pentru evaluarea probabilităţii de blocare (a apelurilor), p_b, nu este necesară rezolvarea sistemului de ecuaţii obţinut deoarece, din punct de vedere al apelurilor, echipamentul pare a fi destinat numai lor, motiv pentru care relaţia de calcul este formula Erlang-*B*, $p_b = B(\rho_1, s)$.

Determinarea soluţiei sistemului de ecuaţii este, totuşi, necesară pentru evaluarea celorlalţi indicatori de performanţă şi se face utilizând aceeaşi metodă ca în cazul anterior. Astfel, se convertesc ecuaţiile III şi IV în 2 ecuaţii algebrice ce conţin, ca necunoscute, funcţiile generatoare:

$$G_0(z) = \sum_{j=0}^{\infty} p_{0,j} \cdot z^j \quad \text{şi} \quad G_1(z) = \sum_{j=0}^{\infty} p_{1,j} \cdot z^j.$$

Ţinând cont de relaţiile de legătură dintre probabilităţi, generate de ecuaţiile I şi II şi aplicând relaţia de normare, $G_0(1) + G_1(1) = 1$, se obţine:

$$p_{0,0} = (1 + \rho_1)^{-1} - \rho_2 \qquad (12.15.9)$$

Pentru ca soluţia să aibă sens, trebuie respectată condiţia:

$$\rho_2 < (1 + \rho_1)^{-1} \qquad (12.15.10)$$

Ţinând cont că $p_b\big|_{M/M/1/\infty/1} = \rho_1/(1+\rho_1)$, această condiţie se poate exprima şi sub forma:

$$\lambda_2 < \mu_2 (1 - p_b) \qquad (12.15.11)$$

ceea ce conduce la următoarele concluzii:

1) Pentru a atinge echilibrul statistic, rata medie de sosire a pachetelor în sistem (λ_2) trebuie să fie mai mică decât capacitatea medie de servire

a pachetelor. Această capacitate este *atenuată* cu factorul $(1-p_b)$ faţă de cazul în care canalul ar fi destinat numai pachetelor. Factorul $(1-p_b)$ reprezintă procentul de timp în care sistemul nu este utilizat de clienţii de tipul 1.

2) Gradul de utilizare a sistemului de către clienţii de tip 2 este limitat la valoarea $1/(1+\rho_1)$ şi nu la 1 ca în situaţia în care pachetele ar fi singurii beneficiari ai sistemului.

Ca şi în cazul integrării fără prioritate, modelul de servire a pachetelor este evaluat prin intermediul indicatorului "media timpului total" petrecut în sistem de un pachet (raportată la timpul de servire), a cărui expresie de calcul este:

$$\mu_2 \cdot \mathrm{E}[T] = \frac{(1+\rho_1)^2 + \alpha\rho_1}{[1-\rho(1+\rho_1)](1+\rho_1)} \qquad (12.15.12)$$

Efectele integrării cu prioritate asupra performanţelor care caracterizează un sistem ce adoptă această soluţie sunt prezentate în continuare pe un caz concret, care presupune că: $s=1$, $1/\mu_1 = 100\,\mathrm{sec}$, $1/\mu_2 = 10\,\mathrm{msec}$, $\rho_1 = 0{,}1\mathrm{E}$ şi $\rho_2 = 0{,}4\mathrm{E}$.

Aplicând formulele de calcul al performanţelor, stabilite anterior, se obţine: $p_b = 0{,}09$ (ca în cazul modelului M/M/1/∞/1) şi $\mu_2 \cdot \mathrm{E}[T_2] = 1600$, ultima valoare evidenţiind creşterea timpului total petrecut în sistem de un pachet, datorată, în principal, scăderii capacităţii canalului ce este oferită pachetelor. În general, considerând $\alpha >> 1$, $\rho_1 << 1$ şi $\rho_2 < 0{,}5$, între cele două strategii de integrare, cu prioritate PA şi respectiv fără prioritate FIFO, se poate face o comparaţie pe baza rapoartelor:

$$\frac{\mathrm{E}[T_2]^{(\mathrm{PA})}}{\mathrm{E}[T_2]^{(\mathrm{FIFO})}} \cong \frac{1}{1-\rho_2} \quad \text{şi} \quad \frac{p_B^{(\mathrm{PA})}}{p_B^{(\mathrm{FIFO})}} \cong \frac{\rho_1}{\rho_2}.$$

12.15.3. Integrarea cu rezervare şi graniţă mobilă

Pentru realizarea integrării cu rezervare, dar folosind o graniţă mobilă se foloseşte următorul procedeu (figura 12.15.5):

- numărul apelurilor din sistem este limitat la o anumită valoarea, s_1;
- graniţa dintre cele N_1 canale alocate apelurilor (clienţi de tipul 1) şi cele $N_2 = s - N_1$ canale alocate pachetelor (clienţi de tip 2) îşi schimbă poziţia în cadru respectând, însă, condiţia $N_1 \le s_1$.

În acest fel, pentru transmiterea pachetelor pot fi puse la dispoziţie canale destinate apelurilor, care temporar nu sunt utilizate. De fapt, această tehnică de integrare este o extensie a integrării cu prioritate absolută, care reprezintă cazul particular: $s_1 = s$ şi $s_2 = 0$.

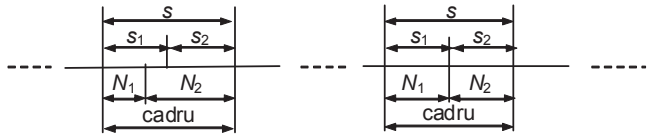


Figura 12.15.5: Principiul rezervării cu graniţă mobilă

Analiza performanţei acestei strategii se face în continuare pe un caz particular ($s=2$ şi $s_1 = s_2 = 1$) şi presupunând că:

- sosirea apelurilor urmează un proces Poisson de rată λ_1;
- sosirea pachetelor urmează un proces Poisson de rată λ_2;
- clienţii ambelor tipuri utilizează un singur canal/cadru;
- timpii de servire sunt distribuiţi exponenţial cu media $1/\mu_1$ pentru apeluri şi $1/\mu_2$ pentru pachete;
- disciplina cozii de pachete este FIFO.

Cum pachetele nu afectează servirea apelurilor pe cele s_1 canale, înseamnă că probabilitatea de blocare a apelurilor este dată de expresia Erlang B, $p_b = \rho_1/(1+\rho_1)$. Pentru ceilalţi indicatori de performanţă, este necesară scrierea sistemului de ecuaţii corespunzător diagramei de tranziţii prezentată în figura 12.15.6.

Aplicaţia 12.15.3

Scrieţi ecuaţiile de echilibru global corespunzătoare diagramei din figura 12.15.6.

Indicaţie: Ecuaţiile se vor numerota cu I, II, III, IV şi V, corespunzător conturului considerat. Ecuaţiile IV şi V conduc la un sistem de două ecuaţii cu 2 necunoscute: $G_0(z)$ şi $G_1(z)$. Pe lângă aceste două necunoscute, ecuaţiile mai conţin şi probabilităţile $p_{1,0}$, $p_{0,1}$ şi, motiv pentru care este necesar ca,

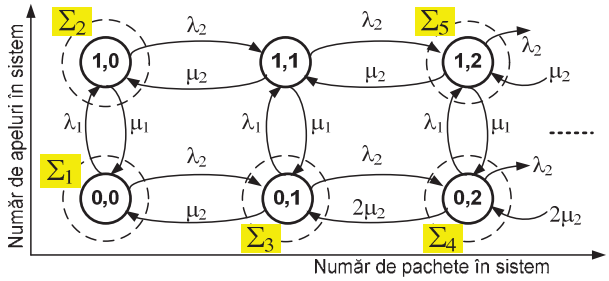


Figura 12.15.6: Diagrama de tranziţii şi stări pentru rezervarea cu graniţă mobilă

înainte de a aplica relaţia de normare, să exprimăm probabilităţile $p_{1,0}$, $p_{0,1}$ în funcţie de probabilitatea $p_{0,0}$.

Acest lucru se face recurgând la ecuaţia numărul I şi la o ecuaţie obţinută prin însumarea termen cu termen a celor două ecuaţii algebrice în care variabilei z i s-a dat valoarea 1. Se ajunge astfel la sistemul de ecuaţii următor:

$$\begin{cases}(\lambda_1+\lambda_2)\cdot p_{0,0}=\mu_2\cdot p_{0,1}+\mu_1\cdot p_{1,0}\\ 2p_{0,0}+p_{0,1}+p_{1,0}=1-\rho_2+(1+\rho_1)^{-1}\end{cases} \quad (12.15.13)$$

Gradul de utilizare a resurselor de către pachete trebuie să respecte relaţia:

$$\left(s-\mathrm{E}[N_1]\right)^{-1}\cdot\lambda_2/\mu_2<1 \quad (12.15.14)$$

care reprezintă condiţia ca rata de sosire a pachetelor să fie mai mică decât rata maximă de servire a lor, termenul $s-\mathrm{E}[N_1]$ fiind media numărului de canale disponibile pentru transmisia pachetelor.

Prin urmare, limita raportului $\rho_2=\lambda_2/\mu_2$ este dată de relaţia:

$$\rho_2<s-\mathrm{E}[N_1]=s-\rho_1(1-p_b)=s_2+\left[s-\rho_1(1-p_b)\right] \quad (12.15.15)$$

în care elementul $\rho_1(1-p_b)$ reprezintă numărul mediu de clienţi corespunzător modelului M/M/s_1/∞/s_1. Gradul total de utilizare este dat de relaţia:

$$\rho=\rho_2+(1-p_b)\cdot\rho_1,$$

care în cazul considerat ($s=2$ şi $s_1=1$) devine:

$$\rho=\rho_2+\rho_1/(1+\rho_1) \quad (12.15.16)$$

Desigur că gradul total de utilizare trebuie să respecte inegalitatea: $\rho<s=2$, pentru ca sistemul analizat să lucreze în zona de echilibru statistic.

Relaţia anterioară o ataşăm sistemului de ecuaţii (12.15.13), ceea ce permite obţinerea următoarei ecuaţii: $2p_{0,0}+p_{0,1}=2-\rho$.

Notând cu r media capacităţii reziduale (neutilizată) a sistemului, a cărei formulă generală de calcul este $r=s-\rho$, iar în cazul particular considerat $r=2-\rho$, rezultă următoarele relaţii de legătură:

$$\begin{aligned}p_{0,1}&=\frac{(2+\rho_1+\alpha_2)\cdot p_{0,0}-r}{\alpha-1}\\ p_{1,0}&=\frac{\alpha r-(\rho_1-\alpha\rho_2+2\alpha)\cdot p_{0,0}}{\alpha-1}\end{aligned} \quad (12.15.17)$$

unde: $\rho_1=\lambda_1/\mu_1$, $\rho_2=\lambda_2/\mu_2$ şi $\alpha=\mu_1/\mu_2$.

Având relaţiile de calcul pentru aceste probabilităţi, transformatele $\mathbf{Z}$ căutate, au expresiile:

$$G_0(z)=\frac{z\cdot(2p_{0,0}+p_{0,1}\cdot z+p_{1,0})+\alpha(z-1)\cdot(1-\rho_2\cdot z)\cdot(2p_{0,0}+p_{1,0}\cdot z)}{\alpha(2-\rho_2\cdot z)\cdot(z-1)\cdot(1+\rho_2\cdot z)+z\cdot(2-\rho_2\cdot z)+\rho_1\cdot z\cdot(1+\rho_2\cdot z)}$$

$$G_1(z)=\frac{\rho_1\cdot z\cdot G_0(z)+\alpha\,p_{1,0}\cdot(z-1)}{\alpha\cdot(z-1)\cdot(1-\rho_2\cdot z)+z}$$

Expresiile obţinute sunt rezultatul unui calcul complex care, din acest motiv, poate fi însoţit de erori. Prin urmare, este util să verificăm dacă formulele obţinute se transformă în relaţii cunoscute atunci când le evaluăm pe cazuri particulare, ce corespund unor modele deja studiate. Astfel, dacă considerăm cazul $\rho_1\to\infty$, adică sistemul serveşte *mai tot timpul* apeluri, înseamnă că probabilităţile ca pachetele să fie servite şi de canalele alocate apelurilor sunt nesemnificative ($p_{0,j}\to 0$ pentru orice j) şi, deci transformatele $\mathbf{Z}$ devin:

$$G_0(z)=\frac{p_{1,0}}{\rho_1(1-\rho_2\cdot z)}\to 0 \quad \text{şi} \quad G_1(z)=\frac{p_{1,0}}{1-\rho_2\cdot z}.$$

În consecinţă, sistemul se comportă din punct de vedere al pachetelor conform modelului M/M/1, acest lucru fiind verificat de expresia lui $G_1(z)$ (vezi funcţia generatoare a probabilităţilor de stare corespunzătoare modelului M/M/1).

În extrema cealaltă, adică dacă $\rho_1\to 0$, ceea ce semnifică o apariţie extrem de rară a apelurilor, probabilităţile stărilor cu un apel în sistem sunt nesemnificative ($p_{1,j}\to 0, \forall j$), transformatele $\mathbf{Z}$ devenind:

$$G_1(z)\to 0 \text{ şi } G_0(z)=p_{0,0}\cdot(2+\rho_2\cdot z)\cdot(2-\rho_2\cdot z)^{-1},$$

ultimul rezultat fiind în concordanţă cu modelul M/M/2 cu care sistemul se confundă în acest caz.

Până în acest punct al raţionamentului, expresiile celor două transformate $\mathbf{Z}$ nu sunt complet determinate, între cele trei probabilităţi, $p_{0,1}$, $p_{1,0}$ şi $p_{0,0}$, identificându-se doar două relaţii de legătură.

A treia ecuaţie se obţine ţinând cont că funcţia $G_0(z)$ reprezintă o funcţie generatoare de momente, care, în consecinţă, trebuie să ia valori finite pentru orice valoare a variabilei z, care respectă condiţia $|z|<1$. Prin urmare, dacă există rădăcini, z_0, ale numitorului care satisfac condiţia $|z_0|<1$, atunci ele trebuie să se regăsească şi la numărător pentru ca efectul lor (de pol al funcţiei) să fie anihilat.

Indicatori de performanţă:

- ***numărul mediu de pachete în sistem:***

$$\begin{aligned}\mathrm{E}[N_2]&=\left.\frac{\mathrm{d}\,G_0(z)}{\mathrm{d}\,z}\right|_{z=1}+\left.\frac{\mathrm{d}\,G_1(z)}{\mathrm{d}\,z}\right|_{z=1}\\ &=\frac{1}{r(1+\rho_1)}\left[\alpha\,p_{1,0}+(1+\rho_1)(\rho_2+p_{0,1})-\frac{\alpha\rho_1}{1+\rho_1}(1-\rho_2)\right]\end{aligned} \quad (12.15.18)$$

Aplicaţia 12.15.5

Determinaţi formele limită ale expresie anterioare în condiţiile: $\rho_1 \to \infty$, şi respectiv $\rho_1 \to 0$. Interpretaţi rezultatele.

Indicaţie: Rezultatele se vor confrunta cu cele obţinute la M/M/2 şi M/M/1.

* * *

➢ ***timpul mediu petrecut de un pachet în sistem*** (raportat la durata de servire). Acest indicator de performanţă are două relaţii de calcul, în funcţie de relaţia dintre parametri ρ_2 şi s_2. Astfel, considerând cazul particular supus analizei, adică $s=2$ şi $s_2=1$, şi presupunând $\alpha >> 1$, ceea ce se respectă în realitate, atunci, pentru cazul $\rho_2 < s_2 = 1$, care corespunde unei funcţionări în condiţii de sarcină normală, se obţine relaţia:

$$\mu_2 \cdot \mathrm{E}[N_2] = \frac{\mathrm{E}[N_2]}{\rho_2} = \frac{4}{r(1+\rho_1)(2+\rho_2)} + \frac{\rho_1}{r(1+\rho_1)} \qquad (12.15.19)$$

iar pentru cazul $\rho_2 > s_2 = 1$, care corespunde unei funcţionări în regim de suprasarcină (în care traficul de pachete depăşeşte capacitatea, s_2, care îi este pusă la dispoziţie fără însă a depăşi limita stabilită în cadrul relaţiei 12.15.15), avem:

$$\mu_2 \cdot \mathrm{E}[N_2] = \frac{1}{r} \cdot \left\{ 1 + \frac{r}{2+\rho_2} + \frac{\alpha(\rho_2 - 1)\rho_1}{\rho_2 (1+\rho_1)^2} \right\} \qquad (12.15.20)$$

Aplicaţia 12.15.6

Verificaţi corectitudinea expresiei timpului mediu, obţinută în cazul funcţionării în regim de sarcină normală.

Indicaţie: Se verifică dacă expresiile obţinute în condiţiile: $\rho_1 \to \infty$, $\rho_1 \to 0$ şi $\rho_2 \to 0$ corespund modelelor de sisteme de servire cu aşteptare corespunzătoare.

* * *

Aplicaţia 12.15.7

Verificaţi continuitatea celor două expresii ale timpului mediu. Interpretaţi rezultatul în cazurile: i) $\rho_1 \to 0$, ii) $\rho_1 \to \infty$.

* * *

Revenind la a doua expresie a timpului mediu petrecut de un pachet în sistem, se observă că este proporţională cu parametrul α care este mult supraunitar ceea ce implică întârzieri mari şi încărcări exagerate ale cozii de aşteptare. Acest lucru nu este de dorit, chiar dacă conduce la o uşoară creştere a sarcinii maxime admise ($\rho_2 < s_2 + \left[s_1 - \rho_1 (1 - p_b) \right]$ în loc de $\rho_2 < s_2$), motiv pentru care, în practică, se utilizează diverse tehnici de control al fluxului de pachete pentru a preîntâmpina funcţionarea în regim de suprasarcină.

Dat fiind complexitatea formulelor de calcul al indicatorilor de performanţă care caracterizează sistemele cu integrarea serviciilor, analizate pe parcursul acestui capitol, (şi nu numai al acestora), o comparaţie sugestivă între soluţii se poate realiza printr-o reprezentare grafică a variaţiilor aceloraşi mărimi. Astfel, figura 12.15.7 compară două strategii de integrare cu graniţă (fixă şi mobilă) din punct de vedere al timpului mediu de aşteptare în coadă, raportat la durata medie de servire a unui pachet. Numărul total de canale considerat este $s=2$ şi alocarea canalelor pe servicii este $s_1 = s_2 = 1$.

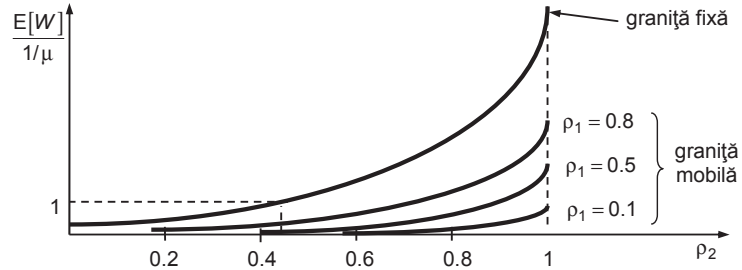


Figura 12.15.7: Comparaţie între diverse strategii de integrare a serviciilor

Se poate observa că strategia cu graniţă mobilă, utilizată în regim normal (sub capacitate $\rho_2 < s_2$), reduce substanţial timpii de aşteptare. Acest lucru se datorează suplimentării cu cantitatea $s_1 - \rho_1(1-p_b)$ a capacităţii destinate pachetelor în cazul sistemului cu graniţă mobilă faţă de sistemul cu graniţă fixă. Surplusul reprezintă media capacităţii (reziduale) neutilizate de apelurile telefonice cărora le sunt dedicate un număr de s_1 canalele. Din acest motiv chiar în condiţiile unui trafic telefonic intens: $\rho_1 = 0{,}8\,\mathrm{E}$ capacitatea disponibilizată de acesta reduce considerabil timpul de întârziere.

Aplicaţia 12.15.8

Demonstraţi următoarea formulă de calcul al timpului mediu de aşteptare care s-a folosit pentru trasarea curbelor din figura de mai sus, corespunzătoare strategiei cu graniţă mobilă: $\mu_2 \cdot \overline{W} = \frac{\rho_2(\rho_2 + 2\rho_1 + \rho_1\rho_2)}{r(1+\rho_1)(2+\rho_2)}$.

12.16 Modele fluide

Multiplexarea este mijlocul prin care (eficienţa utilizării) *gradul de utilizare* al unei resurse, de exemplu procesor sau mediu de conexiune, este îmbunătăţit ca urmare a deservirii simultane a mai multor utilizatori precum fire de execuţie (*thread*), fluxuri informaţionale etc. Modul în care se desfăşoară multiplexarea

depinde de o serie de factori de ordin tehnic şi economic, soluţiile concrete intrând în una din următoarele două categorii:

- **multiplexare deterministă**,
- **multiplexare statistică**.

Concret, în cazul mediilor de comunicaţii, multiplexarea deterministă, întâlnită, de exemplu, în sistemele E1 sau GSM, constă în împărţirea capacităţii (benzii) acestora în canale (circuite) fizice (***channelization***), care sunt puse la dispoziţia utilizatorilor (de exemplu, fluxuri de 64 kb/s) într-o manieră individuală (vezi figura 12.16.1). În schimb, în cazul multiplexării statistice întreaga capacitate a resurselor este utilizată pe rând de fiecare utilizator, conform unei anumite discipline de servire.

Privitor la aceste două tipuri de multiplexare trebuie a fi menţionate următoarele aspecte importante:

- multiplexarea statistică este mult mai eficientă decât multiplexarea deterministă atât timp cât fluxurile individuale ce alcătuiesc fluxul agregat sunt necorelate, cu tot surplusul de informaţie (antet, ...) necesar refacerii la destinaţie a fluxurilor individuale;
- multiplexarea statistică, spre deosebire de multiplexarea deterministă, introduce întârzieri variabile, dat fiind necesitatea transmiterii pe rând a elementelor de flux (pachete, cadre etc.), principiul aplicat fiind **memorează şi înaintează** (*store and forward*);
- ambele metode pot introduce pierderi, dar din cauze diferite:
  - în cazul multiplexării deterministe: debitul unui anumit flux depăşeşte capacitatea canalului repartizat lui,
  - în cazul multiplexării statistice: capacitatea de stocare a elementelor de flux este depăşită.

12.16.1. Modelare matematică

Evaluările de performanţă privind multiplexarea statistică recurg la modelarea analitică sau la simularea pe calculator, cea din urmă fiind însoţită de

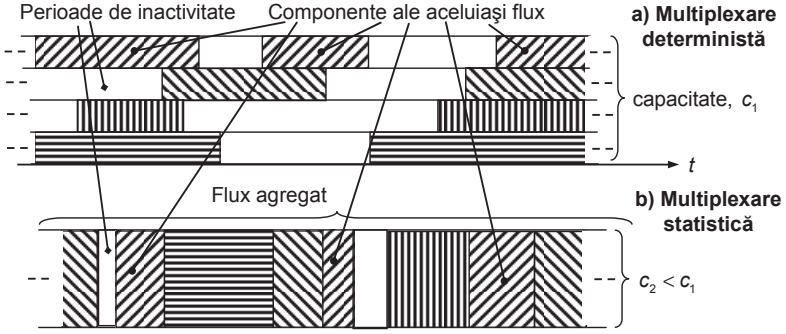


Figura 12.16.1: Tipuri fundamentale de multiplexare

o analiză statistică a rezultatelor. În privinţa modelării analitice aceasta se diferenţiază în funcţie de nivelul de detaliere a fenomenului urmărit, la nivel macroscopic fiind dezvoltate **modele fluide** (***fluid flow model***), iar la nivel microscopic beneficiind de rezultatele oferite de **teoria aşteptării** (*queuing theory*), precum cele prezentate în paragrafele anterioare.

În cazul modelelor fluide, acumularea elementelor de flux se face în timp, la debitul corespunzător, iar, în cazul şirurilor cu aşteptare, acumularea are loc instantaneu, relativ la timpii de servire asociaţi. Drept exemplificare, figura 12.16.2 prezintă modul de variaţie a sarcinii (**workload**) acumulată într-un sistem de servire, cu o capacitate relativă $c = 2h$, considerat din ambele puncte de vedere. În privinţa modelării prin şiruri de aşteptare, s-a considerat că disciplina de servire, adică ordinea în care sunt servite elementele de flux acumulate, este de tip FIFO.

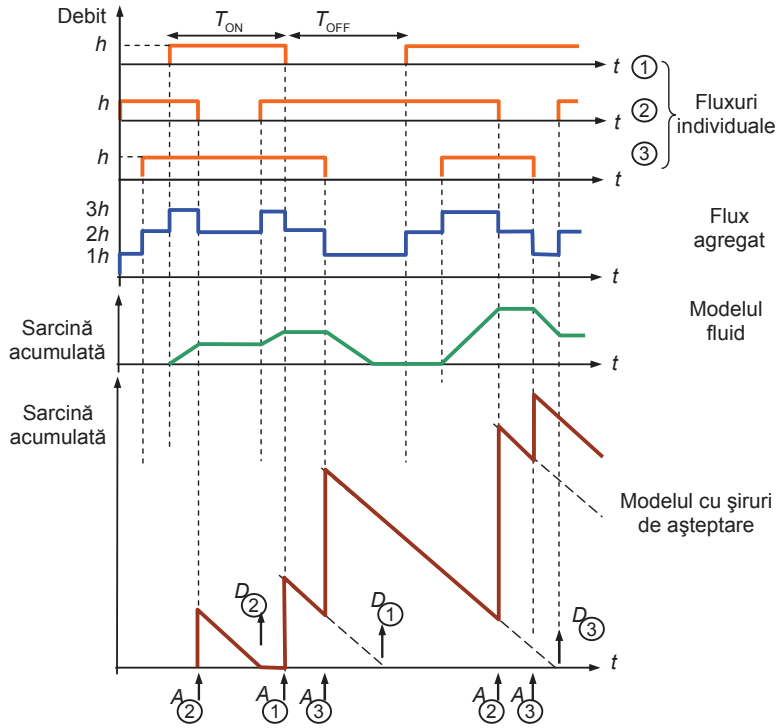


Figura 12.16.2: Variaţia sarcinii acumulate funcţie de modelul considerat

12.16.1.1. Modelul fluid cu acumulare

Modelul fluid [Hay] cu acumulare are ca principală aplicaţie estimarea funcţiei de distribuţie a încărcării unui tampon ajuns la echilibru, la un anumit trafic oferit. Pe baza funcţiei de distribuţie, se poate calcula lungimea medie a tamponului ceea ce permite ca, prin aplicarea formulei Little, să se stabilească întârzierea medie în tampon

12.16.1.1.1. Modelul surselor de trafic

Caracteristica fundamentală a majorităţii claselor de trafic deservite de reţelele/sistemele de comunicaţii este, cum s-a mai menţionat, **intermitenţa**, adică faptul că o sursă de trafic basculează între două stări: de repaus şi de activitate. Drept urmare, modelul analitic, care "mimează" acest comportament, poartă denumirea de surse ON-OFF.

În general, duratele de repaus şi de activitate, T_{OFF} şi T_{ON} din figura 12.16.2, pot urma diverse distribuţii. Atunci când, în ambele cazuri, este vorba de distribuţia exponenţială, sursele ON-OFF în cauză sunt **procese debit cu întrerupere, IRP** (*Interrupted Rate Process*), adică procese cu debit modulat de un proces Markov cu două stări, ca în figura 12.16.3.

În starea OFF, sursa generează informaţie cu debitul $r_0 = 0$, iar în starea ON cu debitul $r_1 = h$ (indicii 0 şi 1 reprezintă corespondenţii numerici ai şirurilor de caractere "ON" şi "OFF", improprii utilizării în cadrul formulelor de calcul).

Parametrii unui proces IRP sunt:

- ***frecvenţa ieşirilor din starea OFF***, λ:

$$\lambda = 1/\mathrm{E}[T_0] \qquad (12.16.1)$$

unde $\mathrm{E}[T_0]$ este durata medie de staţionare în starea 0 (starea OFF)

- ***frecvenţa ieşirilor din starea ON***, μ:

$$\mu = 1/\mathrm{E}[T_1] \qquad (12.16.2)$$

unde $\mathrm{E}[T_1]$ este durata medie de staţionare în starea 1 (starea ON)

- ***debitul de vârf***, h, exprimat în octeţi/sec.

Ţinând cont de aceşti parametri, principalele **caracteristici ale unui proces IRP** au următoarele formule de calcul:

- ***factorul de activitate***, care reprezintă probabilitatea stării ON la echilibru (procentul de timp cât sursa este activă):

$$\alpha = \lambda/(\lambda+\mu) \qquad (12.16.3)$$

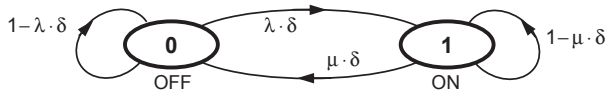


Figura 12.16.3: Diagrama de tranziţii a procesului IRP

- ***lungimea medie a unei rafale:***

$$\mathrm{E}[N_B] = h \cdot 1/\mu \qquad (12.16.4)$$

- ***debitul mediu***:

$$m = h\frac{\lambda}{\lambda+\mu} \qquad (12.16.5)$$

- ***varianţa debitului:***

$$\sigma^2 = h^2\frac{\lambda\cdot\mu}{(\lambda+\mu)^2} \qquad (12.16.6)$$

- ***autocovarianţa***

$$\mathrm{COV}(\tau) = h^2\frac{\lambda\cdot\mu}{(\lambda+\mu)^2}\cdot\exp\left[-(\lambda+\mu)\cdot|\tau|\right] \qquad (12.16.7)$$

Aplicaţia 12.16.1.1

Verificaţi corectitudinea relaţiilor de calcul (12.16.3) - (12.16.7).

Indicaţie: Se notează procesul modulator cu $S(t)$ şi procesul IRP cu $W(t)$. În cazul covarianţei se pleacă de la relaţia generală:

$$\mathrm{COV}(\tau) = \mathrm{E}[\mathrm{W(t)}\cdot\mathrm{W(t+\tau)}] - \mathrm{E}^2[\mathrm{W(t)}] \qquad (12.16.8)$$

Ţinând cont de probabilitatea de legătură (de ordinul 2) a stărilor, primul termen, adică **funcţia de corelaţie statistică**, $R(\tau)$, se poate scrie în forma:

$$\begin{aligned} R(\tau) = \mathrm{E}[\mathrm{W(t)}\cdot\mathrm{W(t+\tau)}] &= \ldots = h^2 P[S(t+\tau)=1, S(t)=1] \\ &= h^2 P\left[S(t+\tau)=1 \,\middle|\, S(t)=1\right]\cdot P[S(t)=1] \end{aligned} \qquad (12.16.9)$$

Dat fiind proprietatea de lipsă a memoriei, caracteristică oricărui proces Markov, înseamnă că:

$$P\left[S(t+\tau)=1 \,\middle|\, S(t)=1\right] = P\left[S(\tau)=1 \,\middle|\, S(0)=1\right]$$

unde termenul din dreapta egalităţii reprezintă probabilitatea stării 1, cu condiţia iniţială $S(0)=1$. Prin urmare, pentru a stabili formula de calcul a probabilităţii căutate, trebuie rezolvat sistemul de ecuaţii ale viitorului, care caracterizează procesul $S(t)$, adică:

$$\mathrm{d}\,\mathbf{p}(t)/\mathrm{d}t = M\cdot\mathbf{p}(t) \qquad (12.16.10)$$

unde $\mathbf{p}(t)$ este vectorul probabilităţii de stare, iar M este matricea generatoare infinitezimală. Folosind, de exemplu, metoda valorilor şi vectorilor proprii (vezi Capitolul 12.16.3), soluţia pentru (12.16.10) este:

$$p_i(t) = k_{i_0} + k_{i_1}\cdot\exp\left[-(\lambda+\mu)\cdot t\right] \text{ cu } i = 0 \text{ sau } 1 \qquad (12.16.11)$$

în care $\left(p_0(t), p_1(t)\right) = \left(P[S(t)=0], P[S(t)=1]\right)$ este vectorul probabilităţilor de stare. Determinarea constantelor se face ţinând seama de condiţiile la limită, care, în cazul stării 1 sunt: $p_1(0)=1$ şi $p_1(\infty)=p_1$, în care p_1 este probabilitatea

staţionară de stare obţinută prin rezolvarea sistemului (12.16.10) în condiţiile $\mathrm{d}\,\mathbf{p}(t)/\mathrm{d}\,t = 0$.

La cele de mai sus, se adaugă faptul că autocorelaţia este o funcţie pară şi, în final, se ajunge la formula (12.16.7).

* * *

Modelarea unui flux agregat, care reprezintă multiplexarea mai multor fluxuri individuale, constă în suprapunerea unui număr corespunzător de surse ON-OFF independente, rezultatul obţinut fiind, în cazul proceselor IRP, aşa zisul **N proces debit cu întrerupere**, **NIRP** (***N** Interrupted **R**ate **P**rocess*), cu diagrama de tranziţii ca în figura 12.16.4.

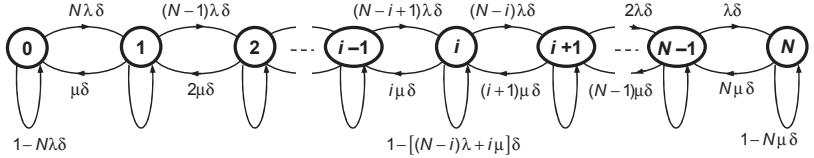


Figura 12.16.4: Diagrama de tranziţii a procesului NIRP

Parametrii unui proces NIRP sunt parametrii proceselor IRP asociate, la care se adaugă numărul N, al acestora.

Ţinând cont de aceşti parametri, principalele caracteristici ale unui **proces NIRP** se calculează astfel:

- **probabilitatea ca un anumit număr de surse să fie active:**

$$p_i = \mathrm{C}_N^i \cdot \alpha^i \cdot (1-\alpha)^{N-i} \qquad (12.16.12)$$

- **debitul mediu:**

$$m_N = N\,h\frac{\lambda}{\lambda+\mu} \qquad (12.16.13)$$

- **varianţa debitului:**

$$\sigma_N^2 = N\,h^2\frac{\lambda\mu}{(\lambda+\mu)^2} \qquad (12.16.14)$$

- **covarianţa debitului:**

$$\mathrm{COV}_{\mathrm{N}}(\tau) = N\,h^2\frac{\lambda\cdot\mu}{(\lambda+\mu)^2}\cdot\exp\left[-(\lambda+\mu)\cdot|\tau|\right] \qquad (12.16.15)$$

12.16.1.2. Modelul fluid fără acumulare

Acest model se pretează în situaţiile în care fluxul agregat provine de la un număr mare de surse de trafic, necorelate, astfel încât teorema limită centrală să poată fi aplicată, iar posibilităţile de stocare ale sistemului sunt nesemnificative, fiind destinate exclusiv rezolvării situaţiilor în care sosirile blocurile informaţionale (precum pachete), din componente distincte ale fluxului,

se suprapun în timp. Prezentarea sa urmează a fi făcută în continuare, prin intermediul aplicaţiei ce urmează.

Aplicaţia 12.16.1.2

Fie un sistem destinat transmisiei printr-un singur canal a fluxurilor numerice provenite de la mai multe surse de trafic. Să se determine economia de bandă ce se obţine dacă sistemul foloseşte multiplexarea statistică în locul celei deterministe, considerând că:

- număr de surse este $n = 100$;
- debitele surselor sunt variabile aleatorii independente, identic distribuite după o lege normală (Gauss) de medie $m\,m$ şi deviaţie standard σ;
- se acceptă o **probabilitate de depăşire a capacităţii** (***overflow***) egală cu $p = 10^{-10}$;
- sistemul nu are posibilităţi de stocare a cererii în exces.

Caz particular: $m = 5\ \mathrm{Mbit/sec}$ şi $\sigma = 1\ \mathrm{Mbit/sec}$.

Rezolvare: Fie $X_1, X_2, \cdots, X_n$ variabile aleatorii independente şi suma lor $S_n = X_1 + X_2 + \ldots + X_n$. Se ştie că funcţia caracteristică a unei variabile aleatorii, X, este definită prin: $\Phi_X(\omega) = \mathrm{E}\left[\mathrm{e}^{j\omega X}\right] = \int_{-\infty}^{+\infty} f_X(x)\cdot\mathrm{e}^{j\omega x}\mathrm{d}x$ putând fi interpretată ca medie a unei funcţii de variabilă aleatorie, $\mathrm{e}^{j\omega X}$, sau ca transformată Fourier (cu semn schimbat la exponent). Relaţia de legătură între funcţia caracteristică a unei sume de variabile aleatorii independente, S_n, şi funcţiile caracteristice ale variabilelor aleatorii individuale este:

$$\Phi_{S_n}(\omega) = \mathrm{E}\left[\mathrm{e}^{j\omega S_n}\right] = \mathrm{E}\left[\mathrm{e}^{j\omega(X_1+X_2+\ldots+X_n)}\right] = \mathrm{E}\left[\mathrm{e}^{j\omega X_1}\cdot\mathrm{e}^{j\omega X_2}\cdot\ldots\cdot\mathrm{e}^{j\omega X_n}\right]$$

$$= \mathrm{E}\left[\mathrm{e}^{j\omega X_1}\right]\cdot\mathrm{E}\left[\mathrm{e}^{j\omega X_2}\right]\cdot\ldots\cdot\mathrm{E}\left[\mathrm{e}^{j\omega X_n}\right] = \Phi_{X_1}(\omega)\cdot\Phi_{X_2}(\omega)\cdot\ldots\cdot\Phi_{X_n}(\omega)$$

În cazul variabilelor aleatorii specificate anterior, caracterizate prin funcţia densitate de probabilitate

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\cdot\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right], \text{ avem:}$$

$\Phi_{X_k}(\omega) = \exp\dfrac{-j\omega m_k + \omega^2\sigma_k^2}{2}$ şi $\Phi_{S_n}(\omega) = \exp\dfrac{-j\omega n\,m + \omega^2 n\sigma_k^2}{2}$, cu $k = \overline{1,n}$.

Rezultă deci că variabila aleatoare S_n urmează tot o distribuţie Gauss, având media $m_\Sigma = n\,m$ şi varianţa $\sigma_\Sigma^2 = n\sigma^2$.

Pentru un debit individual, probabilitatea acceptată de pierdere este dată de relaţia $P[X > x] = p$, iar pentru debitul de grup, de relaţia $P[S_n > x] = p$. Ultima probabilitate se obţine prin amplificarea celei dintâi cu n, considerând că pierderile de debit se distribuie uniform între cele n surse. Valoarea x reprezintă limita maximă, peste care apar pierderi de debite.

În cazul distribuţiei normale, pierderile anterioare pot fi calculate cu ajutorul funcţiei $Q(y)=P[Y>y]$, numită "coada" distribuţiei, a cărei valoare este dată de aria haşurată din figura 2.6.2. Curba trasată în figura amintită reprezintă funcţia densitate de probabilitate a unei variabile aleatoare ce urmează o distribuţie Gauss, de medie nulă şi varianţă egală cu 1.

Valorile funcţiei $Q(y)$ sunt tabelate (vezi Capitolul 2.6.10) şi pot fi folosite pentru orice distribuţie Gauss $f_X(x)$, de medie m şi varianţă σ arbitrare, făcând schimbarea de variabilă: $y=(x-m)/\sigma$. Astfel, debitul maxim acceptat de la o sursă individuală se obţine, în condiţiile impuse de pierderi, din ecuaţia:

$$Q[(x-m)/\sigma]=p=10^{-10}$$

Rezolvarea acestei ecuaţii se poate face prin calcul numeric sau prin consultarea unor tabele special întocmite (tabelul 2.6.2), obţinându-se rezultatul: $(x-m)/\sigma=\alpha=6,36$, sau, altfel spus: $x=m+\alpha\sigma$.

În mod asemănător, se obţine şi debitul maxim acceptat de la toate sursele, plecând de la relaţia: $Q\left[(x_\Sigma-m_\Sigma)/\sigma_\Sigma\right]=p=10^{-10}$, şi folosind aceleaşi tabele: $(x_\Sigma-m_\Sigma)/\sigma_\Sigma=\alpha=6,36$, adică: $x_\Sigma=m_\Sigma+\alpha\cdot\sigma_\Sigma$.

În condiţiile de pierderi impuse, banda necesară transmiterii fluxurilor provenite de la cele n surse este:

- pentru multiplexare deterministă: $B_d=nx=n(m+\alpha\sigma)$
- pentru multiplexare statistică: $B_s=x_\Sigma=nm+\alpha\sigma\sqrt{n}$

ceea ce înseamnă că economia este:

$$B_d-B_s=(\alpha n-\alpha\sqrt{n})\cdot\sigma=(636-63,6)\cdot 1\cong 570\ \text{Mb/s}$$

valoare care evidenţiază avantajul oferit de multiplexarea statistică.

* * *

Metoda folosită în aplicaţia anterioară poate fi utilizată şi în cazul în care debitele surselor individuale urmează alte distribuţii dacă numărul acestora depăşeşte valoarea 20, valoare pentru care, conform teoremei limită centrală, suma unor variabile aleatorii independente tinde către o distribuţie normală.

12.16.2. Studii de caz [Hay]

Aplicaţia 12.16.2.1

Folosind mediul de calcul MatLab şi procesul NIRP drept model de sursă, realizaţi un program care să simuleze o sursă de trafic agregat care incorporează $k=10$ fluxuri (*streams*) video necorelate, care, conform măsurătorilor, prezintă următoarele caracteristici individuale: media $m=3,9$ Mbit/sec şi autocovarianţa $COV(\tau)=3,015\cdot\exp[-3,9|\tau|]\,(\text{Mbit/sec})^2$.

În urma execuţiei, programul va genera un fişier text, intitulat `nirptrace.txt` care conţine o traiectorie a procesului NIRP în cauză.

Verificaţi precizia simulării comparând media şi covarianţa pe eşantion a traiectoriei cu valorile corespunzătoare, caracteristice sursei de trafic real.

Indicaţii:

- Rubrica **SourceModelling**, aferentă **Fişierelor însoţitoare** ale cărţii de faţă, conţine două fişiere utile demersului avut în vedere, [www]. Este vorba de: `SourceModeling.xls` şi `irptrace.m`;
- Parametrii λ, μ şi N ai procesului NIRP se determină în cadrul fişierului `SourceModeling.xls`, conform prezentării teoretice de mai sus şi ţinând cont că, experimental, s-a arătat că precizia unei simulări este suficientă atunci când $N=20k$. Câteva precizări utile, privind folosirea aplicaţiei Excel, se găsesc în fila (*sheet*) `help` a fişierului amintit. Pe lângă acestea, se aminteşte faptul că, pentru a introduce într-o celulă, de exemplu H10, o relaţie de calcul care implică valori din diverse celule, de exemplu: A8+C3, este necesar ca, după ce cursorul devine activ în celula gazdă (H10), să se tasteze simbolul "=", urmând ca apoi să se selecteze celula A8, să se tasteze simbolul "+", să se selecteze celula C3 şi, în final, să se apese tasta "Enter".
- Programul de simulare (vezi aplicaţia 7.1.7) se realizează având ca punct de plecare conţinutul fişierului `irptrace.m`, care, în vederea adaptării la cerinţele temei, se copiază într-un fişier nou, intitulat `nirptrace.m`. Conţinutul nemodificat simulează o sursă IRP de parametrii λ, μ şi h, rezultatele generate fiind:
  - înregistrarea traiectoriei în fişierul `irptrace.txt`
  - calcularea statisticilor,
  - trasarea grafică a traiectoriei şi a autocovarianţei.
- Dat fiind faptul că, în programul ce va fi realizat, evoluţia timpului este discretă, trebuie stabilite condiţiile în care o sursă NIRP discretă aproximează satisfăcător o sursă NIRP continuă. Figura 12.16.6 evidenţiază diferenţele dintre cele două procese, în cazul $N=1$, precizia aproximării crescând odată cu micşorarea intervalului δ.

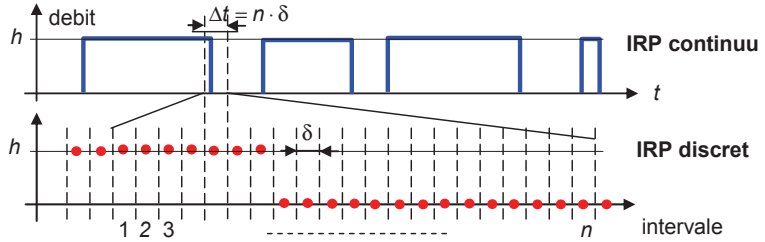


Figura 12.16.6: Traiectorii echivalente a două procese IRP

Ţinând cont de valorile medii ale timpilor de staţionare în stările procesului NIRP, valoarea intervalului δ se poate stabili pe baza relaţiei ce urmează:

$$\delta \cong 1/20 \cdot \min\left(1/(N\lambda), 1/(N\mu)\right) \qquad (12.16.60)$$

- De la un pas la altul al simulării, selectarea tranziţiei se face generând un număr aleatoriu, $0 < U \le 1$ şi verificând în ce segment din intervalul $(0;1]$ este cuprins (vezi figura 12.16.7 şi Capitolul 10);

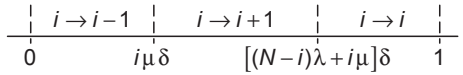


Figura 12.16.7: Scheme de selectare a tranziţiei

- Evitarea regimului tranzitoriu, care poate afecta precizia rezultatelor experimentale, se face stabilind starea iniţială a procesului la o valoare apropiată de numărul mediu teoretic de surse active la un moment dat;
- Rezultatele teoretice se compară cu intervalele de încredere corespunzătoare, obţinute în urma executării programului de mai multe ori (de exemplu: 10).

Aplicaţia 12.17.38

În urma observării portului de ieşire din una din interfeţele de reţea ale unui ruter (la modul abstract, un sisteme de servire), s-a obţinut înregistrarea înscrisă în **fişierul** `"__trace.txt"`, **însoţitor** cărţii de faţă, care precizează, în a doua coloană, numărul pachetelor (clienţi) din port la momentele trecute în prima coloană, corespunzătoare sosirii sau ieşirii clienţilor din sistem.

a) Să se determine:

i) Numărul mediu de clienţi;
ii) Timpul mediu de tranzit;
iii) Timpul mediu de aşteptare;
iv) Intensitatea traficului servit;
v) Intensitatea traficului oferit;
vi) Probabilitatea de respingere;
vii) Probabilitatea de blocare.

Indicaţii: folosind Excel şi ţinând cont de reprezentarea din figura 12.17.4:

i) se calculează aria de sub curba traiectoriei numărului de clienţi sistem şi se împarte la durata de observare;

ii) se calculează rata de tranzit prin sistem și se aplică formula Little;

iii) se calculează timpul mediu de servire;

iv) se calculează probabilitatea ca în sistem să existe cel puţin un client;

v) se calculează produsul dintre rata medie a sosirilor şi timpul mediu de servire;

vi) se numără clienţii respinşi şi se împarte la numărul total de clienţi sosiţi pe durata de observare;

vii) se împarte timpul cât sistemul a fost blocat la timpul total de observare.

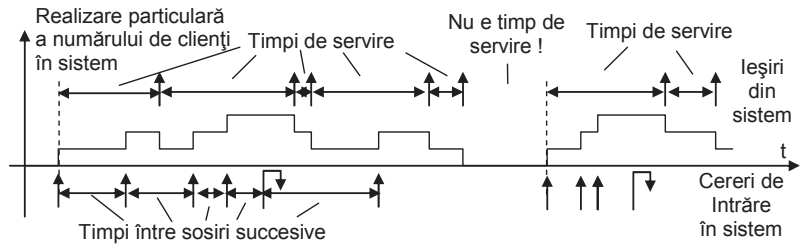


Figura 12.17.4: Sosiri, plecări şi număr de clienţi într-un sistem

b) Să se identifice modelul analitic corespunzător şi să se compare rezultate experimentale cu cele teoretice.

Indicaţie: se determină natura proceselor de sosire şi de servire, adică distribuţiile corespunzătoare şi parametrii asociaţi, pe bază histogramelor şi testelor de conformitate (vezi fişierele `"procesul sosirii... "` şi `"procesul servirii... "`).

Procesul sosirii Sheet 1/3

Stabileste numarul sosirilor in fiecare interval

moment	stare		Se copie doar valorile, fara formule, din coloana C cu "Paste Special" si se sorteaza ascendent	Se copiaza din coloana D doar valorile diferite de zero	Precizeaza capetele intervalelor a cate 10 u.t.	Se aplica instrumentu "Histogram" din meniul "Data Analysis". Pentru Excel 2010, daca meniul nu este instalat, se urmeaza: File>Options>Adds-Ins>Analysis ToolPak, iar pentru versiuni anterioare, se gaseste in "Tools". In rubrica "Input range" se trec datele din coloana E, iar in "Bin Range", cele din coloana F. Rezultatele sunt generate intr-o fila (Sheet) noua.
A	**B**	**C**	**D**	**E**	**F**	**G**
0	0	sosiri (identificate)	sosiri (ordonate)	sosiri (diferite de zero)	capete intervale (bin)	
16.894659	1	16.894659	0	16.894659	10	
17.557159	0	0	0	18.997337	20	
18.997337	1	18.997337	0	20.358044	30	
20.358044	2	20.358044	0	31.01569	40	
23.349838	1	0	0	33.214695	50	
24.927338	0	0	0	33.357323	60	
31.01569	1	31.01569	0	34.119873	70	
33.214695	2	33.214695	0	47.736004	80	
33.357323	3	33.357323	0	58.225346	90	
33.357323	2	0	0	61.353653	100	
34.119873	3	34.119873	0	63.840961	110	
34.119873	2	0	0	64.271172	120	
34.895691	1	0	0	64.543625	130	
37.91069	0	0	0	71.071182	140	
47.736004	1	47.736004	0	82.794609	150	
58.225346	2	58.225346	0	84.845741	160	
61.353653	3	61.353653	0	96.643738	170	

Procesul sosirii Sheet 2/3

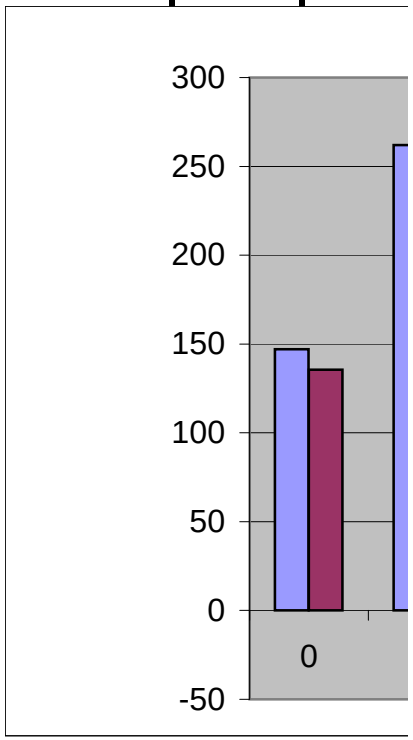
Stabileste numarul de intervale cu acelasi numar de sosiri (frecventa absoluta)

A	B	C	D	E	F	G
Bin	Frequency	Spatiul realizarilor		nr.min sosiri in 10 u.t.=		
10	0	0		nr.max sosiri in 10 u.t.=		
20	2	1				
30	1	2				
40	4	3				
50	1	4				
60	1	5				
70	4	6				
80	1	7				
90	2	8				
100	2					
110	4					
120	3					
130	2					
140	7					
150	0					
160	6					
170	6					
180	4					
190	2					

Procesul sosirii Sheet 3/3

Compara vizual si aplica testul χ^2 pentru verificarea validitatii ipotezelor considerate

A	B	C	D	E	F	G	H	I
Bin	*Frequency*	prob. Poisson	asteptate Poisson	prob. Binomial	asteptate binomial		nr.total intervale=	
0	147	0.135470686	135.4706862				nr. total sosiri=	
1	262	0.270941372	270.9413724				nr med sosiri pe interval=	
2	255	0.270941372	270.9413724					
3	188	0.180627582	180.6275816					
4	104	0.090313791	90.31379081					



Procesul servirii Sheet 1/2

Stabileste numarul sosirilor cu durate cuprinse in intervalele care acopera spatiul in care timpul de servire ia valori

Moment	Stare				Se copie continutul din coloana E cu "Paste Special" si se sorteaza ascendent				marimi caracteristice ale timpului de servire	
A	B	C	D	E	F	G	H	I	j	K
0	0	Serviri	Momentul ultimei serviri	Timpi servire (identificati)	Timpi servire (fara formula si ordonati)	Timpi servire (grupatii)	capete intervale (bin)		maxim=	
16.894659	1	0	16.894659	0	0	0.002441	0.02		minim=	
17.557159	0	1	17.557159	0.6625	0	0.002442	0.04		media=	
18.997337	1	0	18.997337	0	0	0.002442	0.06		std=	
20.358044	2	0	18.997337	0	0	0.004882	0.08			
23.349838	1	1	23.349838	4.352501	0	0.007813	0.1			
24.927338	0	1	24.927338	1.5775	0	0.009766	0.12			
31.01569	1	0	31.01569	0	0	0.012512	0.14			
33.214695	2	0	31.01569	0	0	0.015137	0.16			
33.357323	3	0	31.01569	0	0	0.025391	0.18			
33.357323	2	0	31.01569	0	0	0.027588	0.2			
34.119873	3	0	31.01569	0	0	0.032471	0.22			
34.119873	2	0	31.01569	0	0	0.034912	0.24			
34.895691	1	1	34.895691	3.880001	0	0.037598	0.26			
37.91069	0	1	37.91069	3.014999	0	0.040039	0.28			
47.736004	1	0	47.736004	0	0	0.049804	0.3			
58.225346	2	0	47.736004	0	0	0.052734	0.32			
61.353653	3	0	47.736004	0	0	0.0625	0.34			
61.353653	2	0	47.736004	0	0	0.067383	0.36			
63.840961	3	0	47.736004	0	0	0.067383	0.38			
63.840961	2	0	47.736004	0	0	0.069824	0.4			
63.891006	1	1	63.891006	16.155002	0	0.070068	0.42			

Procesul servirii Sheet 2/2

Aplica testul Kolmogorov-Smirnov de ambele parti, la un nivel probabilistic semnificativ de 0.05

capete intervale	numar serviri cu durate in intervalul corespunzator									
A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**j**	**K**
Bin	*Frequency*		distrib empirica	distrib exp	Dn - exp	distributia uniforma	Dn - unif		dim esantion=	
0	*0*		0	0	0	0	0		deviatia critica=	
0.02	8		0.00538358	0.005115078	0.0002685	0.000625	0.004759		distributia exp Dn max=	
0.04	5		0.008748318	0.010203993	0.00145568	0.00125	0.007498			
0.06	3		0.01076716	0.015266877	0.00449972	0.001875	0.008892		distributia uniforma Dn max=	
0.08	8		0.01615074	0.020303864	0.00415312	0.0025	0.013651			
0.1	11		0.023553163	0.025315086	0.00176192	0.003125	0.020428			
0.12	8		0.028936743	0.030300676	0.00136393	0.00375	0.025187			
0.14	4		0.031628533	0.035260764	0.00363223	0.004375	0.027254			
0.16	5		0.034993271	0.040195481	0.00520221	0.005	0.029993			
0.18	7		0.039703903	0.045104956	0.00540105	0.005625	0.034079			
0.2	4		0.042395693	0.049989319	0.00759363	0.00625	0.036146			
0.22	9		0.048452221	0.054848698	0.00639648	0.006875	0.041577			
0.24	7		0.053162853	0.059683221	0.00652037	0.0075	0.045663			
0.26	4		0.055854643	0.064493015	0.00863837	0.008125	0.04773			
0.28	3		0.057873486	0.069278206	0.01140472	0.00875	0.049123			
0.3	8		0.063257066	0.074038921	0.01078186	0.009375	0.053882			
0.32	7		0.067967699	0.078775285	0.01080759	0.01	0.057968			
0.34	9		0.074024226	0.083487421	0.0094632	0.010625	0.063399			